



AJ

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :  C12Q 1/68 // (C12Q 1/68 C12R 1/32)		A1	(11) International Publication Number: <b>WO 93/22454</b>  (43) International Publication Date: 11 November 1993 (11.11.93)												
<p>(21) International Application Number: <b>PCT/EP93/01063</b></p> <p>(22) International Filing Date: 30 April 1993 (30.04.93)</p> <p>(30) Priority data:</p> <table> <tr> <td>875,940</td> <td>30 April 1992 (30.04.92)</td> <td>US</td> </tr> <tr> <td>929,206</td> <td>14 August 1992 (14.08.92)</td> <td>US</td> </tr> <tr> <td>92/11098</td> <td>17 September 1992 (17.09.92)</td> <td>FR</td> </tr> <tr> <td>93/04545</td> <td>16 April 1993 (16.04.93)</td> <td>FR</td> </tr> </table> <p>(71) Applicants (for all designated States except US): INSTITUT PASTEUR [FR/FR]; 25-28, rue du Dr.-Roux, F-75724 Paris Cédex 15 (FR). MEDICAL RESEARCH COUNCIL [GB/GB]; 20 Park Crescent, London WIN 4AL (GB). ASSISTANCE PUBLIQUE [FR/FR]; 3, avenue Victoria, F-75004 Paris (FR). UNIVERSITE PIERRE ET MARIE CURIE [FR/FR]; 4, place Jussieu, F-75252 Paris Cédex 05 (FR). UNIVERSITE DE BERNE [CH/CH]; Friedbühlstrasse 51, CH-3010 Berne (CH).</p> <p>(72) Inventors; and  (75) Inventors/Applicants (for US only) : HEYM, Beate [DE/FR]; 16, rue de la Fidélité, F-75010 Paris (FR). COLE, Stewart [GB/FR]; 23 bis, rue Cécile-Dinant, F-92140 Clamart (FR). YOUNG, Douglas [GB/GB]; 44 Lawn Close, Ruislip, Middlesex HA4 6ED (GB). ZHANG, Ying [CN/GB]; 137A O'Driscoll House, Ducane Road, London W12 0UE (GB). HONORE, Nadine [FR/FR]; 11, rue des Glycines, F-92700 Colombes (FR). TELENTI, Amalio [ES/CH]; Schlupf 19, CH-3115 Gerzense (CH). BODMER, Thomas [CH/CH]; Sandrüttiweg 9, CH-3423 Ersigen (CH).</p> <p>(74) Agents: GUTMANN, Ernest et al.; 67, boulevard Haussmann, F-75008 Paris (FR).</p> <p>(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p>				875,940	30 April 1992 (30.04.92)	US	929,206	14 August 1992 (14.08.92)	US	92/11098	17 September 1992 (17.09.92)	FR	93/04545	16 April 1993 (16.04.93)	FR
875,940	30 April 1992 (30.04.92)	US													
929,206	14 August 1992 (14.08.92)	US													
92/11098	17 September 1992 (17.09.92)	FR													
93/04545	16 April 1993 (16.04.93)	FR													

(54) Title: RAPID DETECTION OF ANTIBIOTIC RESISTANCE IN MYCOBACTERIUM TUBERCULOSIS

## (57) Abstract

Multidrug resistant strains of *Mycobacterium tuberculosis* represent a considerable threat to public health worldwide. Resistance to isoniazid (INH), rifampicin or analogues thereof, or streptomycin, i.e. key components of anti-tuberculosis regimens, need frequently to be detected. The invention involves the detection of a mutation in either the *katG* gene (isoniazid resistance), the *rpoB* gene (rifampicin resistance) or *rpsL* gene (streptomycin resistance).

## Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LJ	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	MC	Monaco	TG	Togo
DE	Germany	MC	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

RAPID DETECTION OF ANTIBIOTIC RESISTANCE IN MYCOBACTERIUM TUBERCULOSIS

This invention relates to the rapid detection of strains of Mycobacterium tuberculosis that are resistant to antibiotics, particularly isoniazid, rifampicin and streptomycin. More particularly, this invention relates to a method of detecting antibiotic resistance in Mycobacterium tuberculosis, e.g. either as a result of mutations in the relevant genes or by nucleic acid hybridization. This invention also relates to a nucleic acid probe and a kit for carrying out the nucleic acid hybridization. The invention further relates to the chromosomal location of the katG gene and its nucleotide sequence.

BACKGROUND OF THE INVENTION

Despite more than a century of research since the discovery of Mycobacterium tuberculosis, the aetiological agent of tuberculosis, by Robert Koch, this disease remains one of the major causes of human morbidity and mortality. There are an estimated 3 million deaths annually attributable to tuberculosis (Snider, 1989), and although the majority of these are in developing countries, the disease is assuming renewed importance in the West due to the increasing number of homeless people and the impact of the AIDS epidemic (Chaisson et al., 1987; Snider and Roper, 1992).

Isonicotinic acid hydrazide or isoniazid (INH) has been used in the treatment of tuberculosis for the last forty years due to its exquisite potency against the members of the "tuberculosis" groups - Mycobacterium tuberculosis, M. bovis and M. africanum (Middlebrook, 1952; Youatt, 1969). Neither the precise target of the drug, nor its

mode of action, are known, and INH treatment results in the perturbation of several metabolic pathways. There is substantial evidence indicating that INH may act as an antimetabolite of NAD and pyridoxal phosphate (Bekierkunst and Bricker, 1967; Sriprakash and Ramakrishnan, 1970; Winder and Collins, 1968, 1969, 1970), and other data indicating that the drug blocks the synthesis of the mycolic acids, which are responsible for the acid-fast character of mycobacterial cell walls (Winder and Collins 1970; Quemard et al., 1991). Shortly after its introduction, INH-resistant isolates of Mycobacterium tuberculosis emerged and, on characterization, were often found to have lost catalase-peroxidase activity and to show reduced virulence in guinea pigs (Middlebrook et al., 1954; Kubica et al., 1968; Sriprakash and Ramakrishnan, 1970).

Very recently, INH-resistance has acquired new significance owing to a tuberculosis epidemic in the USA due to multidrug resistant (MDR) variants of M. tuberculosis (CDC, 1990; 1991a, b) and the demonstration that such strains were responsible for extensive nosocomial infections of HIV-infected individuals and health care workers (Snider and Roper, 1992). In view of the gravity of this problem, there exists a need in the art to determine the relationship between INH-resistance and catalase-peroxidase production.

More particularly, there is a need in the art to understand the molecular mechanisms involved in drug sensitivity. In addition, there is a need in the art to develop a simple test permitting the rapid identification of INH-resistant strains. Further, there is a need in the art for reagents to carry out such a test.

Rifampicin too is a majeur antibiotic used for the treatment of infections by mycobacterium, particularly Mycobacterium tuberculosis and Mycobacterium leprae. Because some mycobacteria grow slowly, possible rapid and efficient tests for the testing of resistance to rifampicin or analogues thereof must be made available. Likewise the invention aims at a rapid detection of strands of Mycobacterium tuberculosis which are resistant to streptomycin. Because of the development of resistance to streptomycin, the latter antibiotic has been used together with other antibiotics, e.g. isoniazid. Thus adequate treatment of tuberculosis should be preceded by rapid and efficient detection of resistances to the three majeur antibiotics, isoniazid, rifampicin and streptomycin.

#### SUMMARY OF THE INVENTION

Accordingly, this invention aids in fulfilling these needs in the art by providing a process for detecting in vitro the presence of cells of a Mycobacterium tuberculosis resistant to isoniazid and other drugs, such as rifampicin or analogues thereof, and streptomycin.

By analogues of rifampicin, a particularly meant derivatives of 3-formyl-rifamycin, particularly as a result of substitution the rein for the substituent present either in the naphtofuranonyl group or of the site chain at position 7 of the naphtofuranonyl group, or by the introduction or removal of a double band in the lateral chain.

In accordance with the invention, the detection of a resistance to isoniazid involves the detection of one or several mutations in the katG gene of Mycobacterium

tuberculosis, particularly with respect to the nucleotide sequence of that same katG gene in Mycobacterium tuberculosis that are not resistant to isoniazid.

Another process alternative for detecting in vitro the presence of nucleic acids of a Mycobacterium tuberculosis resistant to isoniazid, wherein the process comprises the steps of:

- contacting said nucleic acids previously made accessible to a probe if required under conditions permitting hybridization;
- detecting any probe that had hybridized to said nucleic acids;

wherein said probe comprises a nucleic acid sequence, which is 2.5 kb EcoRV-KpnI fragment of plasmid pYZ56 or of part thereof, and wherein said fragment contains a BamHI cleavage site, wherein said part is nonetheless sufficiently long to provide for the selectivity of the in vitro detection of a Mycobacterium tuberculosis resistant to isoniazid.

For instance, this process alternative comprises the steps of :

- (A) depositing and fixing nucleic acids of the cells on a solid support, so as to make the nucleic acids accessible to a probe;
- (B) contacting the fixed nucleic acids from step (A) with a probe under conditions permitting hybridization;
- (C) washing the filter resulting from step (B), so as to eliminate any non-hybridized probe; and then
- (D) detecting any hybridized probe on the washed filter resulting from step (C).

The probe comprises a nucleic acid sequence which is

present in a 2.5 kb EcoRV-KpnI fragment of plasmid pYZ56, wherein said fragment contains a BamHI cleavage site. This fragment has been found to be associated with intracellular DNA of isoniazid-sensitive Mycobacterium tuberculosis and is capable of distinguishing such antibiotic sensitive microorganisms from isoniazid-resistant Mycobacterium tuberculosis, which do not contain DNA that hybridizes with this fragment under the conditions described hereinafter.

This invention further provides nucleotide sequences, such as RNA and DNA, of isoniazid-resistant Mycobacterium tuberculosis encoding the region of the katG gene of Mycobacterium tuberculosis that imparts isoniazid sensitivity absent from isoniazid-resistant cells.

This invention also provides a probe consisting of a label, such as a radionuclide, bonded to a nucleotide sequence of the invention.

In addition, this invention provides a hybrid duplex molecule consisting essentially of a nucleotide sequence of the invention hydrogen bonded to a nucleotide sequence of complementary base sequence, such as DNA or RNA.

Also, this invention provides a process for selecting a nucleotide sequence coding for a catalase-peroxidase gene of Mycobacterium tuberculosis, or for a portion of such a nucleotide sequence, from a group of nucleotide sequences, which comprises the step of determining which of the nucleotide sequences hybridizes to a nucleotide sequence of the invention. The nucleotide sequence can be a DNA sequence or an RNA sequence. The process can include the step of detecting a label on the nucleotide sequence.

Further, this invention provides a kit for the

detection of Mycobacterium tuberculosis resistant to isoniazid. The kit comprises a container means containing a probe comprising a nucleic acid sequence, which is a 2.5 kb EcoRV-KpnI fragment of plasmid pYZ56, wherein the fragment contains a BamHI cleavage site. The kit also includes a container means containing a control preparation of nucleic acid.

The invention also covers compounds obtained as products of the action of the enzyme catalase, or a similar enzyme on isoniazid. The katG gene or a derivative of this gene which retains a similar activity can be used as a source of catalase protein. The new compounds are selected by reactivity on INH-resistant-mycobacterial strains by the antibiogram method such as described in H. David et al.'s "Méthodes de laboratoire pour Mycobacteriologie clinique" edited by Pasteur Institute, ISBN N° 0995-2454.

#### BRIEF DESCRIPTION OF THE DRAWINGS

This invention will be described in greater detail by reference to the drawings in which:

Fig 1. shows the INH-resistant M. smegmatis strain, BH1 (Gayathri et al., 1975) (a derivative of strain MC<sup>2</sup>-155) was transformed with a pool of M. tuberculosis-H37Rv shuttle cosmids (kindly provided by Dr. W.R. Jacobs, New York) and individual clones were scored for INH-susceptibility. Cosmid pBH4 consistently conferred drug susceptibility and the transformant overproduced catalase (assayed as in Heym, 1992). The restriction map of the DNA insert from pBH4 is shown along with that of the insert from pYZ55 - a plasmid containing katG of M. tuberculosis H37Rv, isolated on the basis of hybridization with an oligonucleotide probe

(5'-TTCATCCGCATGGCCTGGCACGGCGGGCACCTACCGC-3') designed to match the amino acid sequence from a conserved region of E. coli hydroperoxidase I (HPI). Restriction sites for the following enzymes are indicated : B, BamH1; C, Clal; E, EcoRV; H, HindIII, K, Kpn1; M, Smal; N, Not1; R, EcoR1; S, Sac1. Transformation of BH1 with a mycobacterial shuttle plasmid, pBAK14, Zhang et al., 1991, containing the 4.5 kb insert from pYZ55 similarly conferred INH-susceptibility. MIC's are also shown for BH1 transformed with subfragments derived from pYZ55 and inserted into pBAK14 in one (+) or other (-) orientation. The katG gene and the ability to confer INH-susceptibility both mapped to a 2.9 kb EcoRV-Kpn1 fragment (pBAK-KE+).

Fig. 2 shows extracts from M. tuberculosis H37Rv and from E. coli strains transformed with a variety of plasmid constructs that were prepared for activity gel analysis as described previously (Zhang et al., 1991). Non-denaturing gels containing 8% polyacrylamide were stained for catalase (panel A) and peroxidase (panel B) activities as described by Wayne and Diaz (Wayne et al., 1986). Lane 1, M. tuberculosis H37Rv; 2, E. coli UM2 (katE, katG; 3, E. coli UM2/pYZ55; 4, E. coli UM2/pYZ56 (the 2.9 kb EcoRV-Kpn1 fragment in pUC19, corresponding to pBAK-KE+ in Fig. 1); 5, E. coli UM2/pYZ57 (pYZ55 with a BamH1-Kpn1 deletion, corresponding to pBAK-KB+ in Fig. 1). M. tuberculosis catalase and peroxidase activities migrated as two bands under these conditions (lane 1); the same pattern was seen for the recombinant enzyme expressed by pYZ55 (lane 3). pYZ56 (lane 4) expresses a protein of increased molecular weight due to a fusion between katG and lacZ from the vector as shown in panel

C. Panel C also shows partial sequence alignment with E. coli HPI.

Fig. 3 shows an E. coli strain with mutations in both katG and katE (UM2 Mulvey et al., 1988) that was transformed with pUC19 vector alone, pYZ55 expressing M. tuberculosis katG and pYZ56 with high level expression of M. tuberculosis katG. Overnight cultures in Luria-Bertani broth supplemented with appropriate antibiotics were plated out in the presence of varying concentrations of INH and colony forming units were assessed. Results of a representative experiment are shown with error bars indicating the standard deviation observed in triplicate samples. Overexpression of M. tuberculosis katG similarly conferred susceptibility to high concentrations of INH in E. coli UM255 (katG, katE, Mulvey et al., 1988), but had no effect on catalase-positive strains such as E. coli TG1. In some experiments, high concentrations of INH had detectable inhibitory effect on growth of UM2 and UM255, alone, but in all experiments inhibition of pYZ56-transformants was at least 10-100 fold greater than that observed in the corresponding vector controls.

Fig. 4 shows Southern blots prepared using genomic DNA from different M. tuberculosis strains, digested with KpnI, that were probed with (A) katG (the 4.5 kb KpnI fragment), and (B) the SOD gene (1.1 kb EcoR1-KpnI fragment, Zhang et al., 1991). Labelling of probes and processing of blots was performed as described previously (Eiglmeier et al., 1991; Maniatis et al., 1989). Lane 1, H37Rv; 2, strain 12 - MIC 1.6  $\mu$ g/ml INH; 3, B1453 - MIC > 50  $\mu$ g/ml INH (Jackett et al., 1978); 4, strain 24 - MIC > 50  $\mu$ g/ml INH; 5, 79112 - INH-sensitive (Mitchison et

al., 1963); 6, 12646 - INH-sensitive (Mitchison et al., 1963); 7, 79665 - INH-sensitive (Mitchinson et al., 1963). INH susceptibilities were confirmed by inoculation of Lowenstein-Jensen slopes containing differing concentrations of INH.

Fig. 5. Organization of the katG locus. The upper bar corresponds to a stretch of the M. tuberculosis chromosome spanning the katG region and the positions of individual cosmids used to construct the map are shown below together with the original shuttle cosmid pBH4 and pYZ55. The locations of some key restriction sites (B, BamHI; K, KpnI) are shown together with the approximate location of the known genetic markers: fbpB encoding the alpha or 85-B antigen (Matsuo et al., 1988); katG, catalase-peroxidase; LL105, an anonymous  $\lambda$ gt11 clone kindly supplied by A Andersen; MPTR, major polymorphic tandem repeat (Hermans et al., 1992).

Fig. 6. A. Nucleotide sequence of the KpnI fragment bearing katG. This sequence has been deposited in the EMBL data-library under accession number X68081. The deduced protein sequence is shown in the one letter code. B. Alignment of the two copies of the 700 bp direct repeat with identities shown as \* and - denoting pads introduced to optimize the alignment. Numbering refers to the positions in Fig. 2A.

Fig. 7. Distribution of katG in mycobacteria. A. Samples of different bacterial DNAs (1.5  $\mu$ g) were digested with RsrII, separated by agarose gel electrophoresis and stained with ethidium bromide; lanes 1 and 7, size markers; M. leprae; lane 3, M. tuberculosis H37Rv; lane 4, M. gordonaiae; lane 5, M. szulgai; lane 6, M. avium. B. Hybridization of the gel in A, after

10

Southern blotting, with a katG specific probe.

Fig. 8. Primary structure alignment of catalase-peroxidases. The sequences are from M. tuberculosis H37RV, mtkatg; E. coli, eckatg (Triggs-Raine et al., 1988); S. typhimurium, stkatg; B. stearothermophilus, bspera (Loprasert et al., 1988) and yeast cytochrome c peroxidase (ccp; Finzel et al., 1984). The alignment was generated using PILEUP and PRETTY (Devereux et al., 1984) and . denote gaps introduced to maximize the homology. Key residues from the active site and the peroxidase motifs (Welinder, 1991), discussed in the text, are indicated below the consensus.

Fig. 9. Western blot analysis of M. tuberculosis KatG produced in different bacteria. Proteins were separated by SDS-polyacrylamide gel electrophoresis then subjected to immunoblotting, and detection with antiserum raised against BCG, as described in Zhang et al., 1991.

Lane 1, soluble extract of M. tuberculosis H37Rv; lane 2, M. smegmatis MC<sup>2</sup>155 harboring the vector pBAK14; lane 3, MC<sup>2</sup>155 harboring pBAK-KK (katG); lane 4, E. coli UM2 (katE, katG), lane 5, UM2 harboring pYZ55 (katG); lane 6, UM2 harboring pYZ56 (lacZ'::katG).

#### DETAILED DESCRIPTION OF PREFERRED EMBODIMENT

The recent emergence of large numbers of strains of M. tuberculosis showing multidrug resistance in the United States is a most alarming development given the extreme contagiousness of this organism. This danger has been strikingly illustrated by several small tuberculosis epidemics in which a single patient infected with MDR M. tuberculosis has infected both HIV-positive individuals, prison guards and healthy nursing staff (CDC 1990, 1991;

11

Daley et al., 1992; Snider and Roper, 1992). Given the gravity of the current worldwide HIV epidemic, it is conceivable that if AIDS patients in the West, like those in Africa, were to be infected with MDR M. tuberculosis strains (rather than members of the M. avium/M. intracellulare complex) widespread dissemination of the disease would result.

Isoniazid (INH) is a bactericidal drug which is particularly potent against the tuberculosis group of mycobacteria - Mycobacterium tuberculosis, M. bovis, and M. africanum - and, in consequence, it has been particularly effective in the treatment of tuberculosis. Standard anti-tuberculosis regimens generally include INH and rifampicin, often in combination with the weaker drugs, pyrazinamide, ethambutol or streptomycin. Besides its use in therapy INH is also given to close contacts of patients as a prophylactic measure.

INH-resistant mutants of M. tuberculosis, the agent of the human disease, show two levels of resistance: low (1 to 10 µg/ml) and high (10 to 100 µg/ml). INH-resistance is often associated with loss of catalase activity and virulence. Recently, owing to the AIDS epidemic, increased homelessness and declining social conditions, tuberculosis has reemerged as a major public health problem in developed countries, particularly the USA. An alarming feature of the disease today is the emergence of multiple drug-resistant organisms and rapid nosocomial transmission to health care workers and HIV-infected patients. This has prompted CDC to propose new recommendations for the treatment of multiple resistant strains (at least INH and rifampicin) and the prevention of transmission. To obtain fresh insight into the

12

problem of INH-resistance and to develop a rapid diagnostic test the following study was performed.

Clearly, it is essential to understand the mechanisms of resistance to INH and rifampicin, the main anti-tuberculosis agents, as this will allow novel chemotherapeutic strategies to be developed and facilitate the design of new compounds active against MDR strains.

This invention demonstrates that it is the catalase-peroxidase enzyme, HPI, which is the INH target, and it is suggested that this enzyme alone mediates toxicity. Compelling evidence of this conclusion was obtained by expression of the M. tuberculosis katG gene in a catalase-negative mutant of E. coli as this resulted in this bacterium becoming sensitive to INH. Moreover, the isolation of the M. tuberculosis INH-sensitivity gene, katG, is important as it will facilitate the rapid detection of INH-resistant strains by means of hybridization and PCR-based approaches. The high frequency of katG deletions in clinical strains, as shown here, should simplify this procedure.

#### Identification of an M. tuberculosis gene involved in INH-sensitivity

A heterologous approach was employed to isolate M. tuberculosis gene(s) involved in INH-sensitivity. BH1 is a spontaneous mutant of the easily transformable M. smegmatis strain MC<sup>2</sup>155 (Snapper et al., 1990), that is resistant to 512 µg/ml of the INH and lacks catalase-peroxidase activity (Heym et al., 1992). As there is a strict correlation between INH-sensitivity and these enzyme activities, transformation of BH1 with a

plasmid carrying the appropriate gene from M. tuberculosis should lead to their restoration and concomitant INH-sensitivity.

Consequently, DNA was prepared from a pool of M. tuberculosis shuttle cosmids in Escherichia coli and introduced into BH1 by electro-transformation. Over 1000 kanamycin-resistant transformants were then scored for INH-sensitivity, and four clones that failed to grow on medium containing 32 g/ml of INH, the MIC from wild type strain MC<sup>2</sup>155, were obtained.

After re-transformation of BH1, only one of these, pBH4, consistently conferred the INH-sensitive phenotype. Restriction digests with BamHI, KpnI, NotI, Clal and HindIII showed the M. tuberculosis chromosomal DNA carried by pBH4 to be about 30 kb in size. A map produced with the last three enzymes is presented in Fig. 1.

When pBH4 was used as a hybridization probe to detect homologous clones in the library, a further eight shuttle cosmids were isolated. On transformation into BH1, five of these (T35, T646, T673, T79, T556) restored INH-sensitivity, and showed similar restriction profiles to pBH4. In particular, a KpnI fragment of 4.5 kb was present in all cases.

Attempts to subclone individual BamHI fragments did not give rise to transformants capable of complementing the lesion in BH1 suggesting that a BamHI site might be located in the gene of interest. In contrast, pBH5, a derivative of pBH4, was constructed by deletion of EcoRI fragments and this showed that a 7 kb segment was not required for restoration of INH-sensitivity.

Transformants harboring shuttle cosmids that

14

complemented the INH-resistant mutation of BH1 were examined carefully and the MICs for several antibiotics were established. In all cases, the MIC for INH had been reduced from 512 to 8  $\mu$ g/ml, a value lower than that of the sensitive strain MC<sup>2</sup>155 (32  $\mu$ g/ml). This hypersensitive phenotype suggested that the recombinant clones might be overproducing an enzyme capable of enhancing INH-toxicity. Enzymological studies showed that these transformants all produced about 2-fold more peroxidase and catalase than the wild type strain MC<sup>2</sup>155, which is INH-sensitive.

In addition to INH, many MDR-strains of M. tuberculosis are no longer sensitive to rifampicin, streptomycin, ethambutol and pyrazinamide. To examine the possibility that there might be a relationship between resistance to INH and these compounds, the MICs of several drugs for various M. smegmatis strains and their pBH4 transformants were determined, but no differences were found.

#### Cloning the M. tuberculosis catalase gene

A 45-mer oligonucleotide probe was designed based on the primary sequences of highly conserved regions in the catalase-peroxidase enzymes, HPI, of E. coli (Triggs-Raine et al., 1989), and Bacillus stearothermophilus (Loprasert et al., 1988). When genomic blots of M. tuberculosis DNA were probed with this oligonucleotide, specific bands were detected in most cases. As KpnI generated a unique fragment of 4.5 kb that hybridized strongly, this enzyme was used to produce a size selected library in pUC19.

Upon screening with the oligonucleotide probe, an appropriate clone, pYZ55, was obtained. A restriction

15

map of the insert DNA is presented in Fig. 1 where it can be seen that this corresponds exactly to part of pBH4. Independent confirmation was also obtained by cross-hybridization.

By means of various subcloning experiments the smallest fragment expressing M. tuberculosis catalase-peroxidase activity in E. coli was found to be a 2.5 kb EcoRV-KpnI fragment which, as expected, contained a cleavage site for BamHI. Partial DNA sequence analysis showed that the katG gene carried by pYZ56 encodes a catalase-peroxidase enzyme that is highly homologous to the HPI enzymes of E. coli and B. stearothermophilus:

M. tuberculosis      APINSWPDNASLDKARRLLWPSKKYGKKLSWADLIV

E. coli                \*\*\*\*\*V\*\*\*\*\*I\*Q\*\*\*Q\*I\*\*\*\*\*FI

B. stearothermophilus\*\*\*\*\*N\*\*\*\*\*C\*GR\*\*RNT\*T\*\*LGPICS

(Fig. 2; Triggs-Raine et al., 1988); (Loprasert et al., 1988). Identical residues are indicated by \*. HPI activity was detected in both E. coli and M. smegmatis by staining (see below).

#### Catalase-peroxidase involvement in INH-sensitivity

Having cloned the M. tuberculosis katG gene, it was of immediate interest to investigate the genetic basis of the association between catalase-negativity and isoniazid resistance. A series of constructs was established in the shuttle vector pBAK14 and used to transform the INH-resistant M. smegmatis mutant BH1. Only those plasmids carrying a complete katG gene produced HPI and restored INH-sensitivity. The smallest of these, pBAK14, carried a 2.5 kb EcoRV-KpnI fragment thus demonstrating that the 2 kb region upstream of katG was not involved, and that catalase-peroxidase activity alone was

sufficient to render mycobacteria susceptible to INH.

Cell-free extracts were separated by non-denaturating polyacrylamide gel electrophoresis and stained for peroxidase and catalase activity. Under these conditions, the M. tuberculosis enzyme gave two bands of peroxidase activity (lane 1) which comigrated with catalase activity (Heym et al., 1992).

When introduced into E. coli, the katG gene directed the synthesis of the same proteins, whereas pYZ56 produced proteins slightly larger in size. This is due to the construction of an in-frame lacZ::katG gene fusion. Activity stains were also performed with cell extracts of M. smegmatis. The presence of the katG gene from the M. tuberculosis in BH1 led to the production of catalase-peroxidase enzyme, which displayed the same electrophoretic mobility as the enzyme made in M. tuberculosis, or in E. coli, and the native HPI of M. smegmatis.

#### Basis of INH-resistance in M. tuberculosis

It has been known for many years that a subset of INH-resistant strains, particularly those resistant to the highest drug concentrations, are of lower virulence in the guinea pig and devoid of catalase activity. Genomic DNA was prepared from several clinical isolates of M. tuberculosis and analyzed by Southern blotting using the 4.5 kb KpnI fragment as a probe. In two highly resistant strains, B1453 and 24, the catalase gene has been deleted from the chromosome whereas in others (Fig. 3), such as strain 12, showing low level resistance it is still present but not expressed. Additional studies showed that the region immediately prior to katG was highly

prone to rearrangements.

M. tuberculosis HPI renders E. coli sensitive to INH

To determine whether the HPI enzyme of M. tuberculosis could confer INH sensitivity on E. coli, a series of catalase mutants was transformed with pYZ56 and the MICs determined. Wild type strains were not susceptible to INH, but mutants lacking both endogenous catalase activities, but harboring pYZ56, showed growth inhibition when high levels of INH (500 µg/ml) were present, whereas untransformed strains were insensitive.

For purposes of this invention, a plasmid containing the restriction endonuclease map shown in Fig. 1 was deposited in strain with the National Collection of Cultures of Microorganisms (C.N.C.M.) of the Institut Pasteur, in Paris, France on May 18, 1992, under culture collection accession No. I-1209. This plasmid contains the nucleic acid sequence of the invention, namely, the 4.5 kb KpnI-KpnI fragment of plasmid pYZ56 having the BamHI cleavage site in the fragment.

In general, the invention features a method of detecting the presence of isoniazid-resistant Mycobacterium tuberculosis in a sample including providing at least one DNA or RNA probe capable of selectively hybridizing to isoniazid-sensitive Mycobacterium tuberculosis DNA to form detectable complexes. Detection is carried out with a sample under conditions which allow the probe to hybridize to isoniazid-sensitive Mycobacterium tuberculosis DNA present in the sample to form hybrid complexes and detecting the hybrid complexes as an indication of the presence of is niazid-sensitive Mycobacterium

tuberculosis in the sample. (The term "selectively hybridizing", as used herein, refers to a DNA or RNA probe which hybridizes only to isoniazid-sensitive Mycobacterium tuberculosis and not to isoniazid-insensitive Mycobacterium tuberculosis.) The sample can be comprised of the Mycobacterium tuberculosis cells or a portion of the cells or cell contents enriched in Mycobacterium tuberculosis nucleic acids, especially DNA. Hybridization can be carried out using conventional hybridization reagents. The particular hybridization conditions have not been found to be critical to the invention.

More particularly, DNA sequences from Mycobacterium tuberculosis can be analyzed by Southern blotting and hybridization. The techniques used for the present invention are described in Maniatis et al. (1989). DNA fragments can be separated on agarose gels and denatured in situ. The fragments can then be transferred from the gel to a water insoluble solid, porous support, such as a nitrocellulose filter, a nylon membrane, or an activated cellulose paper, where they are immobilized for example, the Hybond<sup>®</sup> membrane commercialized by Amersham can be used. After prehybridization to reduce non-specific hybridization with the probe, the solid support is hybridized to the nucleic acid probe of the invention. The solid support is washed to remove unbound and weakly binding probe, and the resulting hybrid duplex molecule is examined. A convenient alternative approach is to hybridize oligonucleotides to the DNA denatured in the gel.

The amount of labeled probe which is present in the hybridization solution will vary widely, depending upon

the nature of the label, the amount of the labeled probe which can reasonably bind to the filter, and the stringency of the hybridization. Generally, substantial excesses of the probe over stoichiometric will be employed to enhance the rate of binding of the probe to the fixed DNA.

Various degrees of stringency of hybridization can be employed. The more severe the conditions, the greater the complementarity that is required for hybridization between the probe and the polynucleotide for duplex formation. Severity can be controlled by temperature, probe concentration, probe length, ionic strength, time, and the like. Conveniently, the stringency of hybridization is varied by changing the polarity of the reactant solution. Temperatures to be employed can be empirically determined or determined from well known formulas developed for this purpose.

Unlike Southern hybridization where DNA fragments are transferred from an agarose gel to a solid support, the method of the invention can also be carried out by oligonucleotide hybridization in dried agarose gels. In this procedure, the agarose gel is dried and hybridization is carried out *in situ* using an oligonucleotide probe of the invention. This procedure is preferred where speed of detection and sensitivity may be desirable. The procedure can be carried out on agarose gels containing genomic or cloned DNA of Mycobacterium tuberculosis.

In addition, the method of this invention can be carried out by transfer of Mycobacterium tuberculosis DNA from polyacrylamide gels to nylon filters by electroblotting. Electroblotting may be desirable where

time is of the essence, because electroblotting is typically faster than capillary blotting developed to transfer DNA from agarose gels. This method can be carried out in conjunction with UV-crosslinking. The polyacrylamide gel containing the samples to be tested is placed in contact with an appropriately prepared nylon filter. These are then sandwiched into an electroblotting apparatus and the DNA is transferred from the gel onto the filter using electric current. After a buffer rinse, the filter is ready to be prehybridized and hybridized or UV-crosslinked.

The method of the invention can be carried out using the nucleic acid probe of the invention for detecting Mycobacterium tuberculosis resistant to isoniazid. The probe can be detected using conventional techniques.

The method of the invention can also detect point mutations in the KatG gene, as well as a partial deletion of that gene.

The nucleotides of the invention can be used as probes for the detection of a nucleotide sequence in a biological sample of M. tuberculosis. The polynucleotide probe can be labeled with an atom or inorganic radical, most commonly using a radionuclide, but also perhaps with a heavy metal. Radioactive labels include  $^{32}P$ ,  $^3H$ ,  $^{14}C$ , or the like. Any radioactive label can be employed, which provides for an adequate signal and has sufficient half-life. Other labels include ligands that can serve as a specific binding member to a labeled antibody, fluorescers, chemiluminescers, enzymes, antibodies which can serve as a specific binding pair member for a labeled ligand, and the like. The choice of the label will be governed by the effect of the label on the rate of

hybridization and binding of the probe to the DNA or RNA. It will be necessary that the label provide sufficient sensitivity to detect the amount of DNA or RNA available for hybridization.

In preferred embodiments of the invention, the probe is labeled with a radioactive isotope, e.g.,  $^{32}\text{P}$  or  $^{125}\text{I}$ , which can be incorporated into the probe, e.g., by nick-translation.

In other preferred embodiments, the probe is labeled with biotin, which reacts with avidin to which is bonded a chemical entity which, when the avidin is bonded to the biotin, renders the hybrid DNA complex capable of being detected, e.g., a fluorophore, which renders the hybrid DNA complex detectable fluorometrically; an electron-dense compound capable of rendering the hybrid DNA complexes detectable by an electron microscope; an antibody capable of rendering the hybrid DNA complexes immunologically detectable; or one of a catalyst/substrate pair capable of rendering the hybrid DNA complexes enzymatically detectable. Prior to contacting the bacteria with the probe, the M. tuberculosis bacteria can be lysed to release their DNA, which is then denatured and immobilized on an appropriate solid, DNA-binding support, such as a nitrocellulose membrane.

Another detection method, which does not require the labeling of the probe, is the so-called sandwich hybridization technique. In this assay, an unlabeled probe, contained in a single-stranded vector, hybridizes to isoniazid-sensitive Mycobacterium tuberculosis DNA, and a labeled, single-stranded vector, not containing the probe, hybridizes to the probe-containing vector, labeling the whole hybrid complex.

The sequences of the invention were derived by dideoxynucleotide sequencing. The base sequences of the nucleotides are written in the 5'-----> 3' direction. Each of the letters shown is a conventional designation for the following nucleotides:

A	Adenine
G	Guanine
T	Thymine
C	Cytosine.

The nucleotides of the invention can be prepared by the formation of 3'-----> 5' phosphate linkages between nucleoside units using conventional chemical synthesis techniques. For example, the well-known phosphodiester, phosphotriester, and phosphite triester techniques, as well as known modifications of these approaches, can be employed. Deoxyribonucleotides can be prepared with automatic synthesis machines, such as those based on the phosphoramidite approach. Oligo- and polyribonucleotides can also be obtained with the aid of RNA ligase using conventional techniques.

The nucleotides of the invention are in a purified form. For instance, the nucleotides are free of human blood-derived proteins, human serum proteins, viral proteins, nucleotide sequences encoding these proteins, human tissue, and human tissue components. In addition, it is preferred that the nucleotides are free of other nucleic acids, extraneous proteins and lipids, and adventitious microorganisms, such as bacteria and viruses.

This invention of course includes variants of the nucleotide sequences of the invention or serotypic variants of the probes of the invention exhibiting the same selective hybridization properties as the probes

identical herein.

The nucleotide sequences of the present invention can be employed in a DNA amplification process known as the polymerase chain reaction (PCR). See, e.g., Kwok et al. (1987). PCR is advantageous because this technique is rapid.

DNA primer pairs of known sequence positioned 10-300 base pairs apart that are complementary to the plus and minus strands of the DNA to be amplified can be prepared by well known techniques for the synthesis of oligonucleotides. One end of each primer can be extended and modified to create restriction endonuclease sites when the primer is annealed to the PBMC DNA. The PCR reaction mixture can contain the PBMC DNA, the DNA primer pairs, four deoxyribonucleoside triphosphates,  $MgCl_2$ , DNA polymerase, and conventional buffers. The DNA can be amplified for a number of cycles. It is generally possible to increase the sensitivity of detection by using a multiplicity of cycles, each cycle consisting of a short period of denaturation of the PBMC DNA at an elevated temperature, cooling of the reaction mixture, and polymerization with the DNA polymerase.

Amplified sequences can be detected by the use of a technique termed oligomer restriction (OR). Single-strand conformation polymorphism (SSCP) analysis can be used to detect DNA polymorphisms and point mutations in a variety of positions in DNA fragments. See, Saiki et al. (1985); Orita et al. (1989). For example, after amplification, a portion of the PCR reaction mixture can be separated and subjected to hybridization with an end-labeled nucleotide probe, such as a  $^{32}p$  labelled adenosine triphosphate end-labeled

probe. In OR, an end-labeled oligonucleotide prob hybridizes in solution to a region of the amplified sequence and, in the process, reconstitutes a specific endonuclease site. Thus, hybridization of the labeled probe with the amplified katG sequence yields a double-stranded DNA form that is sensitive to selective restriction enzyme digestion. After restriction with an endonuclease, the resulting samples can be analyzed on a polyacrylamide gel, and autoradiograms of the portion of the gel with the diagnostic labeled fragment can be obtained. The appearance of a diagnostic fragment (e.g., 10-15 bases in length) in the autoradiogram indicates the presence of katG sequences in the PBMCs.

Since it may be possible to increase the sensitivity of detection by using RNA instead of chromosomal DNA as the original template, this invention contemplates using RNA sequences that are complementary to the DNA sequences described herein. The RNA can be converted to complementary DNA with reverse transcriptase and then subjected to DNA amplification.

EXPERIMENTAL PROCEDURESBacterial strains and plasmids

Table 1 outlines the properties of the bacterial strains and plasmids used in this invention.

Table 1. Bacterial Strains And Plasmids

<u>Strains/plasmids</u>	<u>Characteristics</u>
<u>E. coli</u> NM554	
<u>E. coli</u> TG1	<u>supE hsd5 thi delta (lac-proAB) [traD36 proAB+ lacI<sup>g</sup> lacZ delta M15]</u>
<u>E. coli</u> UM2	KatE
<u>E. coli</u> UM255	KatE
<u>M. tuberculosis</u> H37Rv	Virulent strain originally isolated from tuberculosis patient
<u>M. tuberculosis</u> 12	Clinical isolate resistant to low levels of INH (1-2 µg/ml)
<u>M. tuberculosis</u> B1453	Clinical isolate resistant to high levels of INH (>50 µg/ml)
<u>M. tuberculosis</u> 24	Clinical isolate resistant to high levels of INH (>50 µg/ml)
<u>M. tuberculosis</u> 79112	Clinical isolate sensitive to INH
<u>M. tuberculosis</u> 12646	Clinical isolate sensitive to INH
<u>M. tuberculosis</u> 79665	Clinical isolate sensitive to INH
<u>M. smegmatis</u> MC <sup>2</sup> 155	MC <sup>2</sup> 6 <u>het</u>
<u>M. smegmatis</u> BH1	MC <sup>2</sup> 155 <u>het katG</u>

Plasmids

pBH4	Shuttle cosmid, <u>katG+</u> , based on pYUB18
pBH5	Deleted version of pBH4, <u>katG+</u> , (7 kb- <u>EcoRI</u> )
pYZ55	pUC19 derivative with 4.5 kb <u>KpnI</u> fragment, <u>kat+</u>
pYZ56	pUC19 derivative with 2.5 kb <u>EcoRV-KpnI</u> fragment ( <u>kat+</u> )
pYZ57	pUC19 derivative with 3.1 kb <u>KpnI-BamHI</u> fragment, <u>kat-</u>
pBAK14	Mycobacterial shuttle vector (Zhang et al., 1991)
pBAK15	Mycobacterial shuttle vector carrying 4.5 kb <u>KpnI</u> fragment ( <u>kat+</u> )
pBAK16	Mycobacterial shuttle vector carrying 2.5 kb <u>EcoRV-KpnI</u> fragment ( <u>kat+</u> )
pBAK17	Mycobacterial shuttle vector carrying 3.1 kb <u>KpnI-BamHI</u> fragment ( <u>kat-</u> )

The M. tuberculosis H37 RV genomic library was constructed in the shuttle cosmid pYUB18 (Snapper et al., 1988) and kindly supplied by Dr. W. R. Jacobs. Other shuttle vectors employed were pYUB12 (Snapper et al., 1988) and pBAK14 (Zhang et al., 1991).

Microbiological techniques and enzymology

Details of antibiotics used, growth conditions, enzymology and MIC determinations can be found in Heym et al., (1992).

Nucleic acid techniques

Standard protocols were used for subcloning, Southern blotting, DNA sequencing, oligonucleotide biosynthesis, etc. (Maniatis et al., 1989; Eiglmeier et al., 1991).

Activity staining

The preparation of cell-free extracts of E. coli and mycobacteria has been described (Heym et al., 1992; Zhang et al., 1991). Native protein samples were separated by polyacrylamide gel electrophoresis as described by Laemmli (1970) except that SDS was omitted from all buffers, samples were not boiled and betamercaptoethanol was not included in the sample buffer. After electrophoresis of 50 - 100 µg protein samples on 7.5% polyacrylamide gels, catalase activity was detected by soaking the gel in 3mM H<sub>2</sub>O<sub>2</sub> for 20 minutes with gentle shaking. An equal volume of 2% ferric chloride and 2% potassium ferricyanide was added and clear bands of catalase activity revealed by illumination with light. Peroxidase activity was detected as brown bands after soaking gels in a solution containing 0.2 - 0.5 mg/ml diaminobenzidine and 1.5 mM H<sub>2</sub>O<sub>2</sub> for 30 - 120 minutes.

To generate a highly toxic compound it seems most likely that the M. tuberculosis HPI enzyme peroxidatively activates INH (Youatt, 1969; Gayathri-Devi et al., 1975). Now that the katG gene has been isolated and characterized, it should be possible to make new derivatives of INH, which can be activated in a similar manner.

#### Example 1

##### Point mutations in the katG gene associated with the isoniazid-resistance of *M. tuberculosis*

It has been shown in a recent study that the catalase-peroxidase of Mycobacterium tuberculosis, encoded by the katG gene, is involved in mediating the toxicity of the potent anti-tuberculosis drug isoniazid or INH. Mutants resistant to clinical levels of INH show reduced catalase-peroxidase activity and, in some cases, this results from the deletion of the katG gene from the chromosome.

Transformation of INH-resistant strains of Mycobacterium smegmatis and M. tuberculosis with the cloned katG gene leads to restoration of drug-sensitivity. Expression of katG in some strains of Escherichia coli renders this naturally resistant organism susceptible to high concentrations of INH.

As some INH-resistant clinical isolates of M. tuberculosis have retained an intact katG gene, the molecular basis of their resistance was investigated. This study was facilitated by the availability of the nucleotide sequence of a 4.7 kb KpnI fragment from the katG region of the chromosome as this allowed primers suitable for PCR analysis to be designed. Eleven pairs of oligonucleotide primers were synthesized (see Table 2) and used to generate PCR-products, of around 280 bp, that covered the complete katG gene and some of the flanking sequences. In control experiments all eleven primer pairs generated PCR products of the expected size, highly suitable for SSCP-analysis, so a panel of 36 INH-resistant strains of M. tuberculosis, of Dutch or French origin, was examined. Many of these strains are multidrug resistant and were isolated from patients who were HIV-seropositive.

Table 2. Sequences of primer pairs used for PCR-SSCP analysis of the *katG* gene of *M. tuberculosis*

Primer Pair #	5'	3'	Length	G+C (%)	Tm	Production
OLIGO1: OLIGO2:	1 GGGGGTTATCGCOGATG GCCCTCGACGGGTATTTC	1765 2052	1782 2034	18 19	66 63	61.8 61.9
OLIGO1: OLIGO2:	2 AACGGCTGTCGGTOGTG GTCGTGGATGGCTAGGTG	2008 2307	2025 2289	18 19	66 63	61.9 61.9
OLIGO1: OLIGO2:	3 TCGACTTGACGCCCTGACG CAGGTCGCGAACATGACAG	2169 2448	2187 2431	19 18	63 66	61.9 61.9
OLIGO1: OLIGO2:	4 OCACAAAGCCAGCTTCGAC GGTTCACGTAGATCAGCCCC	2364 2647	2382 2628	19 20	53 50	61.9 61.9
OLIGO1: OLIGO2:	5 GCAGATGGGCTGATCTACG ACCTCGATGCGCTGGTGT	2622 2909	2641 2892	20 18	60 66	51.9 51.9
OLIGO1: OLIGO2:	6 GCTGGAGCAGATGGCTTG ATCCACCOGGAGGAGAG	2829 3114	2847 3097	19 18	63 66	61.9 61.9
OLIGO1: OLIGO2:	7 GCCACTGACCTCTCGCTG CGCCCATGCGGTGAAAC	3088 3384	3105 3367	18 18	66 66	61.9 61.9

30

Primer Pair #	8	GCGAAGGCAAGATTGCCAGCC	3304	3322	19	63	61.9	285
OLIGO1:		ACAGCCACCCGAGCACGAC	3588	3571	18	66	61.9	
OLIGO2:								
Primer Pair #	9	CAAACGTCTCTTCCCGAOC	3549	3568	20	60	61.9	281
OLIGO1:		CACCTAACAGCACCGTCATC	3829	3810	20	60	61.9	
OLIGO2:								
Primer Pair #	10	TGCTCGACAAACGGAAACCTG	3770	3789	20	60	61.9	280
OLIGO1:		TCCGAGTTGGACCGAAGAC	4049	4030	20	60	61.9	
OLIGO2:								
Primer Pair #	11	TACCAAGGGCAAGGATGGCAG	3973	3992	20	60	61.9	280
OLIGO1:		GCAAAACACCCAGCACCCCG	4252	4235	18	66	61.9	
OLIGO2:		{#courri r10}						

31

Two of them gave no PCR fragment, with any of the primers used, indicating that katG had been deleted. The remaining 34 strains all yielded the expected PCR products and these were analyzed on SSCP gels so that possible point mutations could be detected. In 20 cases, abnormal strand mobility was observed, compared to that of katG from drug-sensitive *M. tuberculosis*, suggesting that mutational events had indeed occurred. The approximate locations of the mutations, as delimited by the PCR primers, are shown in Table 3.

32Tabl 3. Preliminary results of PCT-SSCP analysis of *katG* from *M. tuberculosis* strains  
x denotes altered mobility; del denotes deletion

Nº	Strain	MIC (INH)	1 1765-	2 2008-	3 2169-	4 2364-	5 2628-	6 2622-	7 3097- 3367	X
1/37	9488		1							
2	9577		1							
3	9112		10							
4	9247		1							
5	9200		1							
6	9116		1							
7/31	9106		1							
8	9291		1							
9/10	9412		1							
11/12	9435		1							
13	9428		1							
14	9441		1							
15/16	9444		1							
17/18	9445		1							
19/20	9330		0,2							
21/22	9420		0,2							
23	9262		0,2							
24/38	9523		1							
25	9592		10							
26	9553		10							
27	9485		10							
28	9181		1							
29	9363		1							
30	9465		1							
32	9178		0,2							
33	9468		0,2							
34	9218		0,2							

SUBSTITUTE SHEET

N°	Strain	MIC (INH)	1 1765- 2034	2 2009- 2289	3 2169- 2431	4 2364- 2628	5 2622- 2892	6 2829- 3097	7 3088 3367
33	9468	0,2							
34	9218	0,2							
35	9503	0,2							
39	9582	1							
41	H37RV	-							
42	Abb	-							
43	Mou	-							
44	13632	>20	del	del	del	del	del	del	del
45	12549	>5	del	del	del	del	del	del	del
46	13749	>20							
47	14006	10							
48	13711	>5							
49	13681	>5							
50	14252	>5							

33

On examination of a 200 bp segment of the katG gene from five independent strains (9188, 9106, 9441, 9444, 9363), a single base difference was found. This was the same in all cases, a G to T transversion at position 3360, resulting in the substitution of Arg-461 by Leu. Thus, in addition to inactivation of katG, INH-resistance can stem from mis-sense mutations that result in an altered catalase peroxidase. This mutation may define a site of interaction between the drug and the enzyme. The results of DNA sequence studies with the remaining mutants are eagerly awaited.

Another conclusion that can be drawn from this study concerns the molecular basis of the multidrug resistance associated with various M. tuberculosis strains. The same mutations are found irrespective of whether a given patient is seropositive or seronegative for HIV. For example, strain 9291, isolated from an HIV-seropositive tuberculosis patient, harbors mutations conferring resistance to INH, rifampin and streptomycin in the katG (R461L), rpoB (S425L) and rpsL (K42R) genes, respectively. The same mutations have been found separately, or in combination, in strains from HIV-seronegative individuals. This means that, for the set of strains studied, there is no novel, single mechanism conferring resistance to several drugs, but rather,

multidrug resistance results from the accumulation of mutations in the genes for distinct drug targets.

Example 2

Nucleotide sequence and chromosomal location of the katG locus of *M. tuberculosis*

Bacterial strains, plasmids and growth conditions. The following bacterial strains from our laboratory collections were used in this study: *M. tuberculosis* H37Rv; *M. smegmatis* MC<sup>2</sup>155 (Snapper et al., 1990); *E. coli* K-12 UM2 (katE katG; Mulvey et al., 1988). The recombinant plasmids, pYZ55 (pUC19, katG'), pYZ56 (pUC19, lacZ'::katG) and the shuttle clones, pBH4 (pYUB18, katG') and pBAK-KK- (pBAK14, katG') have been described recently (Zhang et al. 1992, Nature) and the katG locus of *M. tuberculosis* is schematized in Fig. 5. Mycobacteria were grown at 37°C in Middlebrook 7H9 medium, while *E. coli* strains were cultivated in L-broth, with appropriate enrichments and antibiotics.

Nucleic acid techniques. Standard techniques were employed for the preparation, labelling and hybridization of DNA (Eiglmeier et al. 1991; Zhang et al. 1992, Infect. Immun.; Zhang et al. 1992, Nature). A shotgun library of random fragments of pYZ55 was prepared in M13mp18 as described previously (Garnier et al., 1986) and sequenced using the modified dideoxy technique (Biggin et al. 1983). Sequences were compiled and assembled into contigs using SAP, and analyzed with NIP, SIP and PIP (Staden 1987) running on a Vax 3100 workstation. Gap closure was obtained by using synthetic oligonucleotide primers, synthesized on an ABI 381 apparatus, and T7 DNA polymerase (Pharmacia) to obtain sequences directly from pYZ55. To search for related sequences in the GenBank database (release 73.1) the FASTA (Pearson et al. 1988) and BLAST (Altschul et al. 1990)

programs were used. The PROSITE (Bairoch 1992) catalog was screened to detect possible motifs present in protein sequences and alignments were done with the PILEUP and PRETTY modules of the GCG sequence analysis package (Devereux et al. 1984).

Western blotting and catalase-peroxidase activity staining. Immunoblotting of polypeptides resolved by SDS-polyacrylamide gel electrophoresis and detection with polyclonal antibodies (purchased from DAKO) raised against *M. bovis* BCG, were as described (Zhang et al. 1992, Infect. Immun., Nature, Mol. Microbiol.). Procedures for detecting catalase and peroxidase activities have been outlined recently (Heym et al. 1992; Zhang et al. 1992, Nature).

#### RESULTS

Nucleotide sequence of the katG locus of *M. tuberculosis*. In previous studies, the complete katG gene was cloned independently in *E. coli* on a shuttle cosmid, pBH4, and on a 4.5 kb KpnI restriction fragment thus giving rise to pYZ55 (Fig. 5; Zhang et al. 1992, Nature). The structural gene for catalase-peroxidase was subsequently localized to a 2.5 kb EcoRV - KpnI fragment by sub-cloning. To deduce the primary structure of this important enzyme and thereby gain some insight into its putative role in the conversion of INH into a potent anti-tuberculous derivative, the nucleotide sequence of the complete insert from pYZ55 was determined. This was achieved by the modified dideoxy-shotgun cloning procedure (Biggin et al. 1993) and gaps between the contigs were closed by using specific primers.

On inspection of the resultant sequence which is shown in Fig. 6A, the 4.5 kb fragment was found to contain 4795 nucleotides with an overall dG+dC content of 64.4%. When this was analyzed for the presence of open reading frames,

with high coding-probability values, a single candidate was detected and, from its size, composition and location, this was identified as katG. The absence of any additional open reading frames, on either strand of the KpnI fragment, ruled out the possibility that genes other than katG were involved in conferring INH-susceptibility.

Further analysis of the sequence showed katG to be preceded by two copies of a 700 bp direct repeat which were 68% identical, with the longest stretch of identity comprising 58 bp (Fig. 6B). When the databases were screened with this sequence no significant homologies were detected. To test the possibility that it could correspond to a new repetitive element in M. tuberculosis, a 336 bp probe, encompassing the 58 bp repeat, was used to probe a partially-ordered cosmid library. Positive hybridization signals were only obtained from clones that were known to carry katG. Likewise, a single restriction fragment was detected in Southern blots of M. tuberculosis DNA digested with restriction enzymes BamHI, KpnI and RsrII thereby indicating that this repetitive sequence is not dispersed.

Chromosomal location of katG. As part of the M. tuberculosis genome project, most of the genes for which probes are available have been positioned on the contig map. From the series of overlapping cosmids shown in Fig. 5 it can be seen that the markers linked to katG are LL105 and fbbB encoding an anonymous antigen and the putative fibronectin binding protein, or alpha antigen (Matsuo et al. 1988), respectively. None of the known insertion sequences IS6110 and IS1081 (Collins et al. 1991; McAdam et al. 1990; Thierry et al. 1990, J. Clin. Microbiol.; Thierry et al. 1990, Nucleic Acids Res.), map to this area of the chromosome although the region upstream of katG is densely populated

with copies of the major polymorphic tandem repeat, MPTR (Hermans et al. 1992; Zhang and Young 1993).

**Presence of katG homologues in other mycobacteria.** INH is exquisitely potent against members of the tuberculosis complex yet shows little, if any, activity against other mycobacteria. To determine whether genes homologous to katG were present in other mycobacteria Southern blots of DNA digested with RsrII were hybridized with a probe prepared from a 2.5 kb EcoRV-KpnI restriction fragment carrying katG from M. tuberculosis. Under conditions of high stringency good signals were obtained from M. leprae and M. avium (Fig. 7) while barely discernible hybridization was observed with M. gordonaee and M. szulgai. It has been shown recently that katG homologues are also present in M. smegmatis and M. aurum (Heym et al. 1992).

**Predicted properties of catalase-peroxidase from M. tuberculosis.** The primary structure of catalase-peroxidase, deduced from the nucleotide sequence of katG, is shown in Fig. 6. The enzyme is predicted to contain 735 amino acids, and to display a molecular weight of 80,029 daltons. A protein of this size has been observed in M. tuberculosis, and both recombinant M. smegmatis and E. coli (see below).

Primary structures are available for several other bacterial catalase-peroxidases including those from E. coli, salmonella typhimurium and Bacillus stearothermophilus (Loewen et al. 1990; Loprasert et al. 1988; Triggs-Raine et al. 1988) and these have been shown to be distantly related to yeast cytochrome c peroxidase (Welinder 1991). As the crystal structure of the latter has been determined (Finzel et al. 1984) this can be used to interpret the sequences of the bacterial enzymes. The M. tuberculosis enzyme shows 53.3% conservation with the enterobacterial HPI enzymes, and

shares 45.7% identity with the protein from B. stearothermophilus. An alignment of the sequences of these four enzymes is shown in Fig. 8, along with that of yeast cytochrome c peroxidase (Welinder 1991). It is apparent that the NH<sub>2</sub> terminus, which has no counterpart in the yeast enzyme, is the most divergent part suggesting that this domain of the protein can tolerate extensive deviation and is not required for catalysis. Experimental support for this interpretation is provided in the form of a LacZ-KatG fusion protein which contains an additional 40 amino acid residues (Fig. 9, lane 6; Zhang et al. 1992, Nature). Addition of this NH<sub>2</sub>-terminal segment does not noticeably interfere with either the catalase or peroxidase reactions effected by KatG as judged by activity staining (Zhang et al. 1992, Nature).

Bacterial catalase-peroxidases are believed to have evolved by means of a gene duplication event and consist of two modules, both showing homology to the yeast enzyme, fused to a unique NH<sub>2</sub>-terminal sequence of about 50 amino acid residues (Welinder 1991). The M. tuberculosis enzyme conforms to this pattern and when searched for internal homology using SIP (Staden 1987) it was clear that the region between residues 55-422 was related to the carboxy terminal domain, consisting of amino acids 423-735. Only one of the two active site motifs typical of peroxidases, present in the PROSITE catalog (Bairoch 1992) was found when the M. tuberculosis catalase-peroxidase primary structure was screened as there are two deviations from the consensus around His<sup>269</sup> where the second motif should be. (Consensus pattern for peroxidase 1: [DET]-[LIVMT]-x(2)-[LIVM]-[LIVMSTAG]-[SAG]-[LIVMSTAG]-H-[STA]-[LIVMFY]; consensus pattern for peroxidase 2: [SGAT]-x(3)-[LIVMA]-R-[LIVMA]-x-[FW]-H-x-[SAC]; (Bairoch

1992). In addition, a possible ATP-binding motif (G-x-x-x-x-G-K-T) was detected (Balroch 1992) but as this partially overlaps the active site its presence may be purely fortuitous (Fig. 8).

By analogy with yeast cytochrome  $\text{c}$  peroxidase (Welinder 1991), it was possible to predict a number of structurally and catalytically important residues all of which are located in the  $\text{NH}_2$ -terminal repeat.  $\text{His}^{269}$  should serve as the fifth ligand of the heme-iron while  $\text{Asp}^{380}$  should be its hydrogen-bonded partner. Other residues predicted to be involved in active site modulation and  $\text{H}_2\text{O}_2$  binding are  $\text{Arg}^{104}$ ,  $\text{Trp}^{107}$ ,  $\text{His}^{108}$ ,  $\text{Asn}^{138}$ ,  $\text{Thr}^{274}$  and  $\text{His}^{275}$  (Fig. 4). According to Welinder's predictions (Welinder 1991),  $\text{Trp}^{320}$  should be a key residue and be required for forming the protein-radical site (Sivaraja et al. 1989).

**Antibody response to *M. tuberculosis* KatG.** To evaluate the possible value of KatG as an immunogen, Western blots were probed with anti-serum raised against *M. bovis* BCG in rabbits. As shown in Fig. 9, the 80 kD catalase-peroxidase is one of the prominent antigens recognized in cell-free extracts of *M. tuberculosis*, and *M. smegmatis* expressing the cloned *katG* gene (lanes 1, 3). Likewise, on introduction of the gene into *E. coli* significant levels of catalase-peroxidase were produced a striking increase in expression was obtained from the *lacZ*'-*katG* gene fusion which directed the synthesis of an 85 kD fusion protein (Fig. 9, lane 6).

The aim of the present study was to determine the nucleotide sequence of the *katG* gene and to use the information obtained to try and understand how its product mediates the INH-susceptibility of *M. tuberculosis* and, possibly, to explain the apparent instability of the *katG*

region of the genome. Repetitive DNA is often a source of chromosomal rearrangements and analysis of the DNA sequence upstream of katG revealed two copies of a 700 bp direct repeat. Since this element appears to be confined to this locus it is unlikely to serve as a target for an event, such as homologous recombination, which could lead to the deletion of the gene that is observed so frequently (Zhang et al. 1992, *Nature*; Zhang and Young 1993). Likewise, as a 70 kb stretch of the chromosome of M. tuberculosis H37Rv, encompassing katG, is devoid of copies of IS6110 and IS1081, these insertion sequences do not appear to be likely sources of instability. Rather, the presence of a cluster of major polymorphic tandem repeats, MPTR (Fig. 5; Hermans et al. 1992) situated upstream of katG, suggests that this might act as a recombinational hotspot. It may remove both the MPTR cluster and katG (Zhang and Young 1993). The availability of the sequence of the katG region will allow primers suitable for the polymerase chain reaction to be designed and thus facilitate studies aimed at both rapid detection of INH-resistance and understanding the molecular basis of chromosomal instability.

Perhaps the most intriguing feature of the M. tuberculosis catalase-peroxidase is its ability to mediate INH-susceptibility. In our current working hypothesis, the drug interacts with the enzyme and is converted by the peroxidase activity into a toxic derivative which acts at a second, as yet unknown, site (Zhang et al. 1992, *Nature*). Although horse radish peroxidase can effect this reaction (Pearson et al. 1988; Shoeb et al. 1985), and produce hydroxyl and organic free radicals, very few bacteria, including other mycobacteria, are sensitive to INH. This is intriguing as they contain genes homologous to katG (Fig. 7).

One explanation for this could be provided by the fact that most bacterial contain two catalases, one of which is a broad spectrum enzyme endowed with peroxidase activity, and that the second catalase, by preferentially removing  $H_2O_2$ , limits the ability of the catalase-peroxidase to oxidize INH. As M. tuberculosis lacks the latter activity its KatG enzyme can convert INH to the lethal form without competition for the electron acceptor.

Alternatively, there may be some unique features of the M. tuberculosis enzyme which promote toxicity or favor the interaction with the drug. Examination of the primary structures of the bacterial catalase-peroxidases was not instructive in this respect as they all share extensive sequence identities and contain two motifs characteristic of the active sites of peroxidases. Furthermore, it has been shown recently that expression of the E. coli katG gene can partially restore INH-susceptibility to drug-resistant mutants of M. tuberculosis suggesting that the endogenous enzyme may not possess any drug-specific properties (Zhang et al. 1993). Sequence comparison with the cytochrome c peroxidase from yeast has provided important information about the structural and functional organization of the KatG protein and led to the identification of the putatively-important catalytic residues (Fig. 8).

Now that the complete sequence of katG is available it will be possible to test some of these hypotheses by site-directed mutagenesis and to overproduce the enzyme so that detailed analysis of the enzymatic reaction, and its products, can be performed in vitro. Likewise, it should be a relatively simple matter to isolate mutants that have retained enzymatic activity but are unable to bind or oxidize INH. Of particular interest is the repetitive structure of

the enzyme and the prediction that the NH<sub>2</sub>-terminal repeat contains the active site for peroxidases. This raises the possibility that katG genes, mutated, or truncated at the 3'-end, could arise. It is conceivable that their products, lacking the normal COOH-terminus which may be required for subunit-subunit interactions (Welinder 1991), would be unstable but still retain low enzyme activity. They would thus confer an intermediate level of INH-susceptibility, between that of katG<sup>+</sup> strains and mutants completely lacking the gene, as is often observed in clinical settings.

The invention may of course make use of a part of the above described 2.5 kb EcoRV-KpnI fragment, said part being nonetheless sufficiently long to provide for the selectivity of the in vitro detection of a *Mycobacterium tuberculosis* resistant to isoniazid.

The invention also relates to a kit for detecting multidrug resistant variants of M. tuberculosis wherein the kit comprises:

- (a) a container means containing a probe for the gene encoding drug resistance; and
- (b) a container means containing a control preparation of nucleic acid.

Needless to say that use can be made of any detection method alternative bringing into play the nucleic sequence specific of nucleic acids of a *Mycobacterium* resistant to isoniazid, e.g. a method using an amplification technique and primers, whereby said primers may either be contained within said specific nucleotidic sequence, in order to provide for amplification fragments containing at least a part of the nucleotide sequence of the above mentioned probe, nonetheless sufficiently long to provide for the selectivity of the in vitro detection of a *Mycobacterium tuberculosis* resistant to

isoniazid, and finally detecting a possible mutation in any of the amplified sequences.

A preferred process alternative (oligotyping) for the detection of resistance to the selected antibiotic comprises:

- fragmenting the relevant gene or part thereof likely to carry the mutation into a plurality of fragments, such as by digestion of said relevant gene by selected restriction enzymes,
- hybridizing these fragments to complementary oligonucleotide probes, preferably a series of labelled probes recognizing under stringent conditions, all of the parts of the relevant gene of a corresponding control DNA of a strain non-resistant to the corresponding antibiotic,
- and relating the absence of hybridization of at least one of said oligonucleotide probes to any of the DNA fragments of the relevant gene of the mycobacterium under study as evidence of the presence of a mutation and, possibly, of a resistance to the corresponding antibiotic, particularly as compared to the running of the test under the same conditions with the same oligonucleotides on the relevant gene(s) obtained from a strain (strains) not resistant to said antibiotic.

Another process alternative (SSCP analysis, i.e. analysis of Single Stranded Conformation Polymorphisms) comprises:

- digesting the DNA to be analyzed, particularly of the relevant gene,
- amplifying the fragments obtained, e.g. by PCR,
- recovering the amplified fragments, and
- separating them from one another according to sizes, e.g. by causing them to migrate, for instance on an electrophoretic gel,
- comparing the sizes of the different fragments with those

obtained from the DNA(s) of one or several control strains not resistant to the antibiotic, which had been subjected to a similar assay, and

- relating the polymorphism possibly detected to the existence of a mutation in the relevant gene, accordingly to a possible resistance to the corresponding antibiotic of the strain from which the DNA under study had been obtained.

Needless to say that any other method, including classical sequencing techniques, can be resorted to for the achievement of the same purpose.

This method includes that known under the expression "oligotyping" for the detection of polymorphisms, reference is advantageously made to the method disclosed by Orita et al. (reference was already made thereto herebefore) for the detection of polymorphisms based on the conformation of single strands.

The relevant gene in the case of resistance to isoniazid is of course the katG gene or a fragment thereof.

In the case of resistance to rifampicin, the relevant gene happens to be the rpoB gene which codes for the  $\beta$  sub-unit of the RNA polymerases of said mycobacteria, or when only part of that gene is being used, preferably that part which includes the codons 400 to 450 of that rpoB gene.

Finally, in the case of resistance to streptomycin, the relevant gene contemplated is that of the rpsL gene that codes for the S12 protein of the small ribosome sub-unit or, when only part of said fragment is being used, preferably that part which includes the codon at the 43 position.

A preferred procedure, particularly in relation to the process alternative making use of PCR amplification is disclosed hereafter.

DNA is obtained from a biological sample (e.g. blood or

sputum) after removal of the cellular debris and lysis of the bacterial cells with an appropriate lysis buffer. PCR amplification can be carried out by classical methods, using a pair of primers, whose sequences are respectively complementary to fragments of each of the strands of the DNA to be amplified.

The procedure may be run further as follows:

- the amplification products (comprising e.g. from 100 to 300 nucleotides) are digested by means of suitable restriction endonuclease,
- the ADN strands obtained from the amplification medium are subjected to denaturation,
- the monostranded DNA strands are deposited on a neutral 5% polyacrylamid gel,
- the monostranded DNA strands are caused to migrate on said gel by means of electrophoresis,
- the DNA fragments that migrated on the polyacrilamid gel are transferred onto a nylon membrane according to a usual electrophoretic blotting technique and hybridized to labelled probes, for instance  $^{32}P$  labelled probes, and
- the migration distances of the DNA fragments subjected to analysis are compared to those obtained from controls obtained under the same conditions of amplification, digestion, denaturation electrophoresis and transfer onto a nylon membrane, whereby said DNA had been obtained from an identical bacterial strain yet sensitive to the antibiotic under study.

For the production of the PCR primers as well as of the polygonucleotides probes used in the above disclosed "oligotyping" procedures, use is advantageously made of those complementary to the rpoB gene of wild M. tuberculosis isolated in a plasmid deposited under number I-12167 at the

CNCM on September 15, 1992.

The invention also relates more particularly to the nucleotidic sequence of a fragment of rpsL gene of Mycobacterium tuberculosis coding for the S12 protein of the small ribosome sub-unit, as well as to the nucleotidic sequence of a mutated rpsL gene fragment deemed responsible of the resistance to streptomycin.

By amplification of that nucleotidic sequence, the nucleotide sequence of the full rpsL gene can be obtained.

Further illustration of the invention will be provided in the following description of additional examples, having regard to the drawings in which:

- figure 10 represents diagrammatically the PCR strategy used for the study of different M.Leprae isolates, showing the coding sequence of rpoB sequence, whereby the sequenced regions are shown by hatched parts, and the position and reference of the amplification primers used being indicated on the upper line, whereas the sequencing primers are indicated bellow it;
- figure 11 represents (A) the nucleotidic sequence of a short region of rpoB carrying the mutations that confer resistance to rifampicin with an indication of the changes of bases in the corresponding alleles and (B) a comparison between the aminoacids sequences of the domain I of region II of the  $\beta$ -sub-unit of the RNA polymerase of E.coli and M.Leprae, whereby the numbers of the residues and the differences in the mutated aminoacids have been indicated; the mutated aminoacid residues associated with rifampicin resistance as well as the frequency of its occurrences have been represented too,
- figure 12 shows a complete sequence of the rpoB gene of M.Leprae,

- figure 13 represents the sequence of part of the rpoB gene of M.tuberculosis,
- figure 14 represents the sequence of a part of the rpsL gene of M.tuberculosis; both the sequence of the full rpsL gene of M.Leprae and that of its expression product, that is the S12 protein (whose starting aminoacid is noted by 1) are indicated. The positions of the ML51 and ML52 primers, as well as of the sequences of part of the rpsL gene of M.tuberculosis are provided belows those of M.Leprae. Only those positions which are different and the corresponding aminoacid changes are indicated.
- figure 15 represents the wild DNA sequence of the rpsL gene fragment coding for the S12 protein of the small ribosome sub-unit that is responsible for the resistance to streptomycin, as well as the corresponding aminoacid sequence of the S12 protein.

Example relative modification in the detection of rifampicin resistance of

The sensitivity to rifampicin has been determined in mice as disclosed by Grosset et al. (and Int. J. Lepr. 57:607-614). The cells of M.Leprae were obtained from mouse paws according to classic procedures. All resistant strains were able to grow in mice which received daily doses of 20 mg/Kg of rifampicin, whereas sensitive strains were killed at low rifampicin concentrations, less than 2 mg/Kg.

Relevant regions of the rpoB gene of extracted DNA was initiated upon using two pairs of biotinylated primers, whose sequences appear in the following table.

49

## TABLE

	<u>Primer</u>	<u>Séquence</u>
10	B <sub>rpo</sub> 22	CAGGACGTCGAGGCGATCAC
	r <sub>p</sub> o23	AACGACGACGTGGCCAGCGT
15	B <sub>rpo</sub> 24	CAGACGGTGTATGGGCGA
	r <sub>p</sub> o25	TCGGAGAAACGAAACGCTC
20	r <sub>p</sub> o32	TCCTCGTCAGCGGTCAAGTA
	r <sub>p</sub> o33	CTTCCCTATGATGACTG
25	r <sub>p</sub> o34	GGTGATCTGCTCACTGG
	r <sub>p</sub> o35	GCCGCAGACGCTGATCA
30	r <sub>p</sub> o36	TTGACCGCTGACGAGGA
	r <sub>p</sub> o37	GCCAGCGTCGATGGCCG
35	-----	

SUBSTITUTE SHEET

Upon using conventional techniques, amplification products comprising 310 and 710 bp were respectively obtained as shown in figure 1. The localization of the sequences of the different primers used in the table is also indicated on figure 10.

The DNAs obtained have been sequenced on the basis of the rpoB sequence of isolates sensitive to rifampicin. A plasmid containing the sequence of that gene has been deposited at CNCM on September 15, 1992 under number I-1266. Biotinylated PCR products were concentrated from the PCR reaction mixtures by contacting with streptavidin coated beads under agitation. The biotinylated strands attached to the beads were then recovered and sequenced. The sequences obtained were compared to the sequence of the rpoB gene of a wild type stain. Significant results were obtained as a result of sequencing of the wild gene (of a mycobacterium sensitive to rifampicin) and of corresponding sequences of the B-sub-unit of four mutant strains resistant to rifampicin (figure 11).

Results were obtained starting from 102 strands obtained from patients infected with M.tuberculosis. Among this 102 strands 53 were sensitive to rifampicin and 49 resistant to rifampicin. The mutation was localized in the region 400-450 in 43 of the mutants and among the latter, the mutation occurred in the region of <sup>425</sup>Ser into leu.

Example of detection of the resistance of mycobacteria to streptomycin

The culture of M.tuberculosis strains and the test of their sensitivity to streptomycin have been carried out by the method of proportions on a Löwenstein-Jerva medium (Laboratory Method for Clinical Mycobacteriology - Hugo David - Véronique Lévy Frébault, M.F. Th rel, published by Institut

Pasteur).

The nucleotide sequence of the rpsL gene of M. Leprae led, by sequence analogy, to the construction of two primers, ML51 (CCCACCATTCAAGCAGCTGGT) et ML52 (GTCGAGCGAACCGCGAATGA) surrounding regions including putative mutation sites liable of being responsible for the streptomycin resistance and suitable for the PCR reaction. The DNA of the used M. tuberculosis used as a matrix has enabled one to obtain a rpsL fragment of 306 pb. The nucleotide sequence of the sequenced fragments exhibited 28 differences with that of M. Leprae.

The rpsL genes of 43 strands of M. tuberculosis, among which 28 were resistant, have been amplified both by PCR and the SSCP technique.

DNA was extracted from 200  $\mu$ l aliquots of M. tuberculosis samples (in average  $10^4$  to  $10^5$  bacteria) covered by 100  $\mu$ l of mineral oil by a congelation-decongelation technique (Woods and Cole, 1989 FEBS. Microbiol. Lett, 65:305-308).

After electrophoresis of the DNA strands tested a mutation was shown in 16 of the mutants. In order to establish the nature of the mutation in the 16 strands under consideration, the corresponding rpsL gene fragments were amplified by PCR using the ML51 and the ML52 primers and their respective nucleotide sequences were determined.

The sequences obtained were compared to the sequence of the wild type rpsL gene. The single difference was found with the wild sequence : codon 43, AAG, was mutated into AGG and, consequently, the lys-42 aminoacid was replaced by Arg.

The invention relates also to the "mutated" DNA fragments. They can in turn be used as hybridization probes for use for the detection in suitable hybridization

procedures and for the detection of similar mutation in DNA extracted from a M. tuberculosis strain suspected to include resistance to any one of the above illustrated antibiotics.

The invention further relates to kits for the resistance of mycobacteriae to isoniazid, rifampicin or analogues thereof, and streptomycin.

The invention further relates to a kit for the in vitro diagnostic of the resistance of a bacteria of a mycobacterium genus to isoniazid, characterized in that it comprises:

- means for carrying out for a genic amplification of the DNA of the katG gene or of a fragment thereof,
- means to bring into evidence one or several mutations on the amplification products so obtained,
- a preparation of control DNA of a katG gene of a strain of said bacteria sensitive to isoniazid or of a fragment thereof,
- optionally, a control preparation of a DNA of the katG gene of an isoniazid-resistant mycobacterium strain.

The invention further relates to a kit for the in vitro diagnostic of the resistance of a bacteria of a mycobacterium genus to rifampicin or its analogues, characterized in that it comprises:

- means for carrying out for a genic amplification of the DNA of the rpoB gene or of the  $\beta$ -sub-unit of the RNA polymerase of said mycobacteria, or of a fragment thereof,
- means to bring into evidence one or several mutations on the amplification products so obtained,
- a preparation of control DNA of a rpoB gene coding for the  $\beta$ -sub-unit of the RNA polymerase of a strain of said bacteria sensitive to rifampicin or of a fragment thereof,
- optionally, a control preparation of a DNA of the rpoB gene of an isoniazid-resistant mycobacterium strain.

Similarly, the invention pertains to a kit for the in vitro diagnostics of the resistance of the M.tuberculosis to streptomycin, characterized in that it includes:

- means for carrying out a genic amplification of the rpsL gene coding for the S12 protein of the small ribosome sub-unit, or fragment thereof,
- means which enable the bringing to evidence of one or several mutations on the amplification products obtained,
- a control preparation of a DNA sequence of the rpsL gene coding for the S12 protein of the small sub-unit of the ribosome of a M.tuberculosis strain sensitive to streptomycin, and
- optionally, a control preparation of a DNA sequence of a rpsL gene coding for the S12 protein of the small sub-unit of the ribosome of a strain of M.tuberculosis resistant to streptomycin.

REFERENCES CITED IN THE SPECIFICATION

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990). A basic local alignment search tool. Proc. Natl. Acad. Sci. USA 87:403-410.

Bekierkunst, A. & Bricker, A. (1967). Studies on the mode of action of isoniazid on mycobacteria. Arch. Biochem. Biophys. 122:385-392.

Biggin, M.D., Gibson T.J., and Hong G.F. (1983). Buffer gradient gels and  $^{35}\text{S}$ -label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA 80:3963-3965.

Bairoch, A., (1992). Prosite: a dictionary of sites and patterns in proteins. Nucleic Acids Res. 20:2013-2018.

C.D.C. Outbreak of multidrug-resistant tuberculosis - Texas, California, and Pennsylvania. MMWR 1990, 39:369-372.

C.D.C. Nosocomial transmission of multidrug-resistant tuberculosis among HIV-infected persons - Florida and New York 1988-1991. MMWR 1991(a) 40:585-591..

C.D.C. Transmission of multidrug-resistant tuberculosis from an HIV-positive client in a residential substance abuse treatment facility. Michigan. MMWR 1991(b), 40:129-131.

Chaisson, R.E., Schechter, G.F., Theuer, C.P., Rutherford, G.W., Echenberg, D.F., Hopewell, P.C. (1987). Tuberculosis in patients with the acquired immunodeficiency syndrome. Am. Rev. Respir. Dis., 23:56-74.

Collins, D.M., and Stephens, D.M. (1991). Identification of an insertion sequence, 1S1081, in Mycobacterium bovis. FEMS Microbiol. Lett. 83:11-16.

Daley, C.L., Small, P.M., Schechter, G.F., Schoolnik, G.K., McAdam, R.A., Jacobs, W.R., and Hopewell, P.C. (1992). An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus.

An analysis using restriction-fragment-length-polymorphism.  
N. Engl. J. Med., 326:231-235.

Devereux, J., Haeberli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. 12:387-395.

Eiglmeier, K., Honore, N., and Cole, S.T. (1991). Towards the integration of foreign DNA into the chromosome of Mycobacterium leprae. Research in Microbiology, 142:617-622.

Finzel, B.C., Poulos, T.L. and Kraut, J. (1984). Crystal structure of yeast cytochrome C peroxidase at 1.7 Å resolution. J. Biol. Chem. 259:13027-13036.

Garnier, T., and Cole, S.T., (1986). Characterization of a bacteriocinogenic plasmid from Clostridium perfringens and molecular genetic analysis of the bacteriocin-encoding gene. J. Bacteriol., 168:1189-1196.

Gayathri Devi, B., Shaila, M.S., Ramakrishnan, T., and Gopinathan, K.P. (1975). The purification and properties of peroxidase in Mycobacterium tuberculosis H37RV and its possible role in the mechanism of action of isonicotinic acid hydrazide. Biochem. J., 149:187-197.

Hermans, P.W.M., van Soolingen, D. and van Embden, J.D.A. (1992). Characterization of a major polymorphic tandem repeat in Mycobacterium tuberculosis and its potential use in the epidemiology of Mycobacterium kansasii and Mycobacterium Aviordonae. J. Bacteriol. 174:4157-4165.

Heym, B. and Cole, S.T. (1992). Isolation and characterization of isoniazid-resistant mutants of Mycobacterium smegmatis and M. aurum. Res. Microbiol., submitted.

Jackett, P.S., Aber, V. and Lowrie, D. (1978). J. Gen Microbiol., 104:37-45.

Kubica, G.P., Jones Jr., W.D., Abbott, V.D., Beam, R.E.,

Kilburn, J.O., and Cater Jr., J.C. (1966). Differential identification of mycobacteria. I. Tests on catalase activity. Am. Rev. Resp. Dis., 94:400-405.

Kwok et al., S., J. Virol. 61:1690-1694 (1987). Multidrug resistance results from the accumulation of mutations in the genes for distinct drug targets.

Laemmli, U.K., (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage-T4. Nature (London) 227:680-685.

Loewen, P.C., and Stauffer, G.V. (1990). Nucleotide sequence of katG of Salmonella typhimurium LT2 and characterization of its product, hydroperoxidase I. Mol. Gen. Genet. 224:147-151.

Loprasert, S., Negoro, S. and Okada, H. (1988). Thermo-stable peroxidase from Bacillus stearothermophilus. J. Gen. Microbiol., 134:1971-1976.

Loprasert, S., Negoro, S., and Okada, H. (1989). Cloning, nucleotide sequence, and expression in Escherichia coli of the Bacillus stearothermophilus peroxidase gene (perA). J. Bacteriol., 171:4871-4875.

Maniatis, T., Sambrook, J., and Fritsch, E.F. (1989). Molecular cloning. A laboratory manual. Second Edition 1989. Cold Spring Harbor Laboratory Press.

Matsuo, K., Yamaguchi, R., Yamazaki, R.A., Tasaka, H. and Yamada, T. (1988). Cloning and expression of the Mycobacterium bovis BCG gene for extracellular  $\alpha$  antigen. J. Bacteriol., 170:3847-3854.

Middlebrook, G. (1954). Isoniazid-resistance and catalase activity of tubercle bacilli. Am. Rev. Tuberc., 69:471-472.

Middlebrook, G., Cohn, M.L., and Schaefer, W.B. (1954). - Studies on isoniazid and tubercle bacilli. III. The

isolation, drug-susceptibility, and catalase-testing of tubercle bacilli from isoniazid-treated patients. Am. Rev. Tuberc., 70:852-872.

Mitchison, D.A., Selkon, J.B. and Lloyd, S. (1963). J. Path. Bact. 86:377-386.

Mulvey, M.R., Sorby PA, Triggs-Raine BL and Loewen PC. Gene 73:337-345 (1988).

Orita, M., Iwahana, I., Kanazawa, H., Itayashi, K., and Sekiya, J. (1989). PNAS 86:2766-2770.

Pearson, W., and Lipman, D. (1988). Improved tools for biological sequence comparisons. Proc. Natl. Acad. Sic. USA. 85:2444-2448.

Quemard, A., Lacave, C., and Laneelle, G. (1991). Isoniazid inhibition of mycolic acid synthesis by cell extracts of sensitive and resistant strains of Mycobacterium aurum. Antimicrob. Ag. Chem., 35:1035-1039.

Saiki et al., R. K., Bio/Technology 3:1008-1012 (1985).

Shoeb, H.A., Bowman B.U.J., Ottolenghi, A.C., and Merola, A.J. (1985). Peroxidase-mediated oxidation of isoniazid. Antimicrobial Agents and Chemotherapy, 27:399-403

Shoeb, H.A., Bowman, B.U.J., Ottolenghi, A.C., and Merola, A.S. (1985). Evidence for the generation of active oxygen by isoniazid treatment of extracts of Mycobacterium tuberculosis H37Ra. Antimicrobial Agents and Chemotherapy, 27:404-407.

Sivaraja, M., Goodin, D.B., Smith, M., and Hoffman, B.M., (1989). Identification by ENDOR of Trp<sup>191</sup> as the free-radical site in cytochrome c peroxidase Compound Es. Science, 245:738-740.

Snapper, S.B., Lugosi, L., Jekkel, A., Melton, R.E., Kieser, T., Bloom, B.R., and Jacobs, W.R. (1988). Lysogeny and transformation in mycobacteria: stable expression of

foreign genes. Proc. Natl. Acad. Sci. USA, 85:6987-6991.

Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., and Jacobs, W.R. (1990). Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis. Mol. Microbiol., 4:1911-1919.

Snider, D. (1989). Rev. Inf. Dis., S335.

Snider Jr., D.E. and Roper, W.L. (1992). The new tuberculosis. The New England Journal of Medicine, 326:703-705.

Sripakash, K.S. and Ramakrishnan, T. (1970). Isoniazid-resistant mutants of Mycobacterium tuberculosis H37Rv: Uptake of isoniazid and the properties of NADase inhibitor. J. Gen. Microbiol., 60:125-132.

Staden, R. (1987). Computer handling of sequence projects. In Nucleic acid and protein sequence analysis: A practical approach. Bishop, M.J. and Rawlings, C.J. (eds.) Oxford: IRL Press, pp. 173-217.

Thierry, D., Brisson-Noël, A., Vincént-Levy-Frébault, V., Nguyen, S., Guesdon, J., and Gicquel, B. (1990). Characterization of a Mycobacterium tuberculosis insertion sequence, IS6110, and its application in diagnosis. S. Clin. Microbiol., 28:2668-2673.

Thierry, D., Cave, M.D., Eisenach, K.D., Crawford, S.T., Bates, S.H., Gicquel, B., and Guesdon, J.L. (1990). IS6110, an IS-like element of Mycobacterium tuberculosis complex. Nucleic Acids Res., 18:188.

Triggs-Raine, B.L., Doble, B.W., Mulvey, M.R., Sorby, P.A., and Loewen, P.C. (1988). Nucleotide sequence of kagG, encoding catalase HPI of Escherichia coli. J. Bacteriol., 170:4415-4419.

Wayne, L.G. and Diaz, G.A. (1986). Analyt. Biochem. 157:89-92.

Welinder, K.G. (1991). Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. Biochim. Biophys. Acta 1080:215-220.

Winder, F. and Collins, P. (1968). The effect of isoniazid on nicotinamide nucleotide levels in Mycobacterium bovis, strain BCG. Amer. Rev. Respir. Dis., 97:719-720.

Winder, F. and Collins, P. (1969). The effect of isoniazid on nicotinamide nucleotide concentrations in tubercle bacilli. Amer. Rev. Respir. Dis., 100:101-103.

Winder, F. and Collins, P. (1968). Inhibition by isoniazid of synthesis of mycolic acids in Mycobacterium tuberculosis, J. Gen. Microbiol., 63:41-48.

Youatt, J. (1969). A review of the action of isoniazid. Am. Rev. Respir. Dis., 99:729-749.

Zhang, Y., Garbe, T., and Young, D. (1993). Transformation with katG restores isoniazid-sensitivity in Mycobacterium tuberculosis isolates resistant to a range of drug concentrations. Mol. Microbiol., submitted.

Zhang, Y., and Young, D.B. (1993) Characterization of a variable genetic element from the katG region of Mycobacterium tuberculosis - in preparation.

Zhang, Y., Lathigra, R., Garbe, T., Catty, D., and Young, D. (1991) Genetic analysis of superoxide dismutase, the 23 kilodalton antigen of Mycobacterium tuberculosis. Mol. Microbiol., 5:381-391.

Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S.T. (1992). The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis. Nature. 358:591-593.

Zhang, Y., Garcia, M.J., Lathigra, R., Allen, B., Moreno, C., van Embden, D.A., and Young, D. (1992). Alterations in the superoxid dismutase gene of an isoniazid-resistant

60

strain of Mycobacterium tuberculosis. Infect. Immun.,  
60:2160-2165.

**SUBSTITUTE SHEET**

CLAIMSWHAT IS CLAIMED IS:

1. A process for the detection of a resistance to an antibiotic in a mycobacterium which comprises detecting a mutation in a gene selected from the group comprising the katG gene or fragment thereof, the rpoB gene or fragment thereof and the rpsL gene or fragment thereof.

2. A process of claim 1 for detecting in vitro the presence of nucleic acids of a Mycobacterium tuberculosis resistant to isoniazid, wherein the process comprises the steps of:

- contacting said nucleic acids previously made accessible to a probe if required under conditions permitting hybridization;
- detecting any probe that had hybridized to said nucleic acids;

wherein said probe comprises a nucleic acid sequence, which is 2.5 kb EcoRV-KpnI fragment of plasmid pYZ56 or of part thereof, and wherein said fragment contains a BamHI cleavage site, wherein said part is nonetheless sufficiently long to provide for the selectivity of the in vitro detection of a Mycobacterium tuberculosis resistant to isoniazid.

3. The process as claimed in claim 2, which comprises the steps of :

- (A) depositing and fixing nucleic acids of Mycobacterium tuberculosis on a solid support, so as to make the nucleic acids accessible to a probe;
- (B) contacting said fixed nucleic acids from step (A) with the probe under conditions permitting hybridization;
- (C) washing said filter resulting from step (B), so as to eliminate any non-hybridized probe; and then

(D) detecting any hybridized probe on said washed filter resulting from step (C);

4. The process of claim 2 or 3 wherein said probe comprises a nucleic acid sequence which encodes a polypeptide of the formula APINSWPDNASLDKAR-RLLWPSKKYGKKLISWADLIV.

5. A process as claimed in any of claims 2 to 4, wherein the probe has a radioactive label selected from the group consisting of radioactive, enzymatic, fluorescent, and luminescent labels.

6. The use of the process of any one of claims 2 to 5 for the detection of the presence of Mycobacterium tuberculosis resistant to isoniazid in a bacteria-containing sample suspected of containing Mycobacterium tuberculosis resistant to isoniazid, whereby the detection of the probe that had hybridized, particularly in the form of a hybrid DNA complex that it either forms or had formed with DNA initially present in said sample, is indicative of the presence in said sample of Mycobacterium tuberculosis resistant to isoniazid.

7. The use of claim 6, wherein prior to the contacting of said DNA with said probe, said bacteria had been separated from said sample and immobilized on a DNA binding support, such as a nitrocellulose membrane.

8. A kit for the detection of Mycobacterium tuberculosis resistant to isoniazid, wherein the kit comprises:

(A) a container means containing a probe, preferably labelled by a label selected from the group consisting of radioactive, enzymatic, fluorescent, and luminescent labels, comprising a nucleic acid sequence, which is a 2.5 kb EcoRV-KpnI fragment of plasmid pYZ56 or part thereof, wherein said fragment contains a BamHI cleavage site and

nonetheless sufficiently long to provide for the selectivity of the in vitro detection of a *Mycobacterium tuberculosis* resistant to isoniazid; and

(B) a container means containing a control preparation of nucleic acid.

9. A nucleic acid probe for detecting *Mycobacterium tuberculosis* resistant to isoniazid, wherein said probe consists of a 2.5 kb EcoRV-KpnI fragment of plasmid pYZ56, wherein said fragment contains a BamHI cleavage site, or of a part of said fragment nonetheless sufficiently long to provide for the selectivity of the in vitro detection of a *Mycobacterium tuberculosis* resistant to isoniazid.

10. The probe as claimed in claim 9, which is DNA free of human serum proteins or human tissue or both, viral proteins, bacterial proteins, and nucleotide sequences encoding said proteins.

11. A hybrid duplex molecule consisting essentially of the probe of claim 9 hydrogen bonded to a nucleotide sequence of complementary base sequence.

12. A process for selecting a nucleotide sequence of a *Mycobacterium tuberculosis* resistant to isoniazid from a group of nucleotide sequences, comprising the step of determining which of said nucleotide sequences hybridizes to a probe as claimed in claim 9 or 10.

13. A process for selecting a compound active against *Mycobacterium tuberculosis* comprising the step of determining the reactivity of the compound on INH-resistant *Mycobacterial* strains.

14. A nucleotide sequence comprising the 350 base sequence or a portion thereof as described in Figure 2.

15. A process for detecting point mutations or partial

deletion of the KatG gene comprising contacting a sample of Mycobacterium tuberculosis with the probe of claim 9 or 10.

16. The process of claim 1 for the detection of resistance to the selected antibiotic which comprises:

- fragmenting the relevant gene or part thereof likely to carry the mutation into a plurality of fragments, such as by digestion of said relevant gene by selected restriction enzymes,

- hybridizing these fragments to complementary oligonucleotide probes, preferably a series of labelled probes recognizing under stringent conditions, all of the parts of the relevant gene of a corresponding control DNA of a strain non-resistant to the corresponding antibiotic,

- and relating the absence of hybridization of at least one of said oligonucleotide probes to any of the DNA fragments of the relevant gene of the mycobacterium under study as evidence of the presence of a mutation and, possibly, of a resistance to the corresponding antibiotic, particularly as compared to results obtained upon running the test under the same conditions with the same oligonucleotides on the relevant gene(s) obtained from a strain (strains) not resistant to said antibiotic, wherein said relevant gene is either the katG gene or a fragment thereof, the rboB gene or a fragment thereof, the rpsL gene or a fragment thereof.

17. The process of claim 1 which comprises:

- digesting the DNA to be analyzed, particularly of the relevant gene,

- amplifying the fragments obtained, e.g. by PCR,

- recovering the amplified fragments, and

- separating them from one another according to sizes, e.g. by causing them to migrate, for instance on an electrophoretic gel,

- comparing the sizes of the different fragments with those obtained from the DNA(s) of one or several control strains not resistant to the antibiotic, which had been subjected to a similar assay, and
- relating the polymorphism possibly detected to the existence of a mutation in the relevant gene, accordingly to a possible resistance to the corresponding antibiotic of the strain from which the DNA under study had been obtained, wherein said relevant gene is either the katG gene or a fragment thereof, the rboB gene or a fragment thereof, the rpsL gene or a fragment thereof.

18. A kit for the in vitro diagnostic of the resistance of a bacteria of a mycobacterium genus to isoniazid, characterized in that it comprises:

- means for carrying out for a genic amplification of the DNA of the katG gene or of a fragment thereof,
- means to bring into evidence one or several mutations on the amplification products so obtained,
- a preparation of control DNA of a katG gene of a strain of said bacteria sensitive to isoniazid or of a fragment thereof,
- optionally, a control preparation of a DNA of the katG gene of an isoniazid-resistant mycobacterium strain.

19. A kit for the in vitro diagnostic of the resistance of a bacteria of a mycobacterium genus to rifampicin or its analogues, characterized in that it comprises:

- means for carrying out for a genic amplification of the DNA of the rpoB gene or of the  $\beta$ -sub-unit of the RNA polymerase of said mycobacteria, or of a fragment thereof,
- means to bring into evidence one or several mutations on the amplification products so obtained,
- a preparation of control DNA of a rpoB gene coding for the

$\beta$ -sub-unit of the RNA polymerase of a strain of said bacteria sensitive to rifampicin or of a fragment thereof,

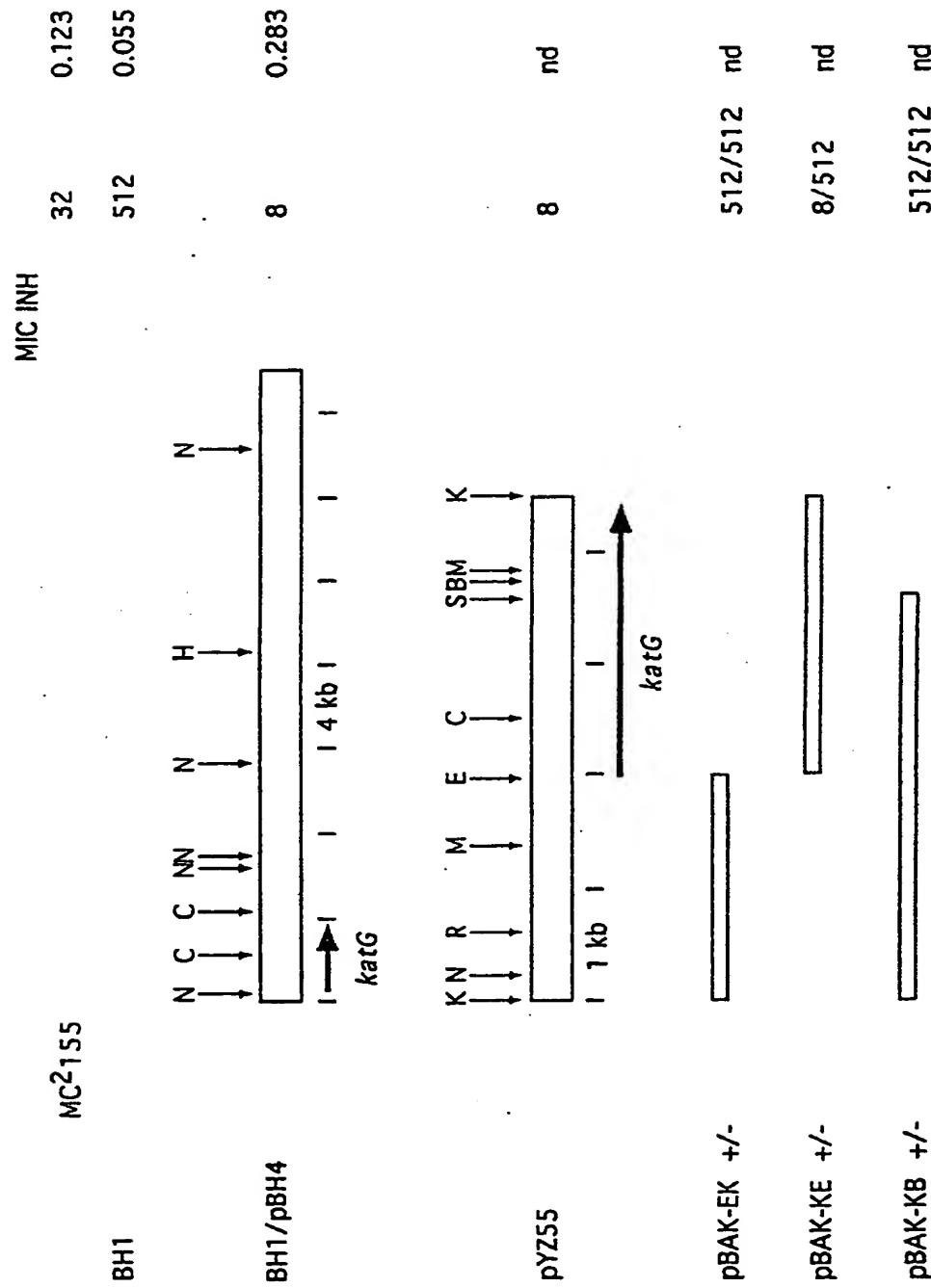
- optionally, a control preparation of a DNA of the rpoB gene of an isoniazid-resistant mycobacterium strain.

20. A kit for the in vitro diagnostics of the resistance of the M. tuberculosis to streptomycin, characterized in that it includes:

- means for carrying out a genic amplification of the rpsL gene coding for the S12 protein of the small ribosome sub-unit, or fragment thereof,
- means which enable the bringing to evidence of one or several mutations on the amplification products obtained,
- a control preparation of a DNA sequence of the rpsL gene coding for the S12 protein of the small sub-unit of the ribosome of a M. tuberculosis strain sensitive to streptomycin, and
- optionally, a control preparation of a DNA sequence of a rpsL gene coding for the S12 protein of the small sub-unit of the ribosome of a strain of M. tuberculosis resis tant to streptomycin.

1 / 26

FIG. 1



SUBSTITUTE SHEET

2/26

FIG. 2A

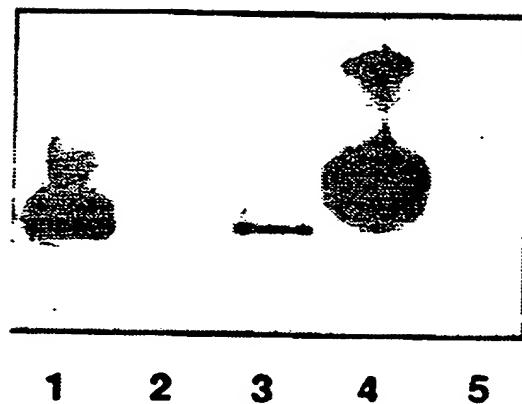
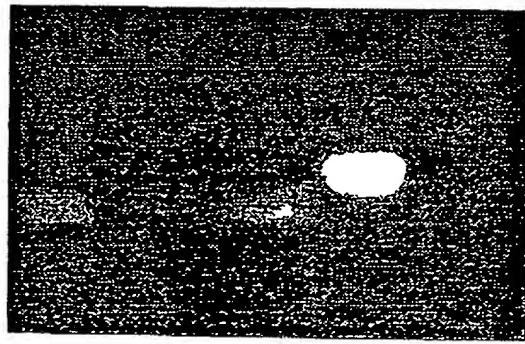


FIG. 2B

**SUBSTITUTE SHEET**

3 / 26

FIG. 2C(1)

## **SUBSTITUTE SHEET**

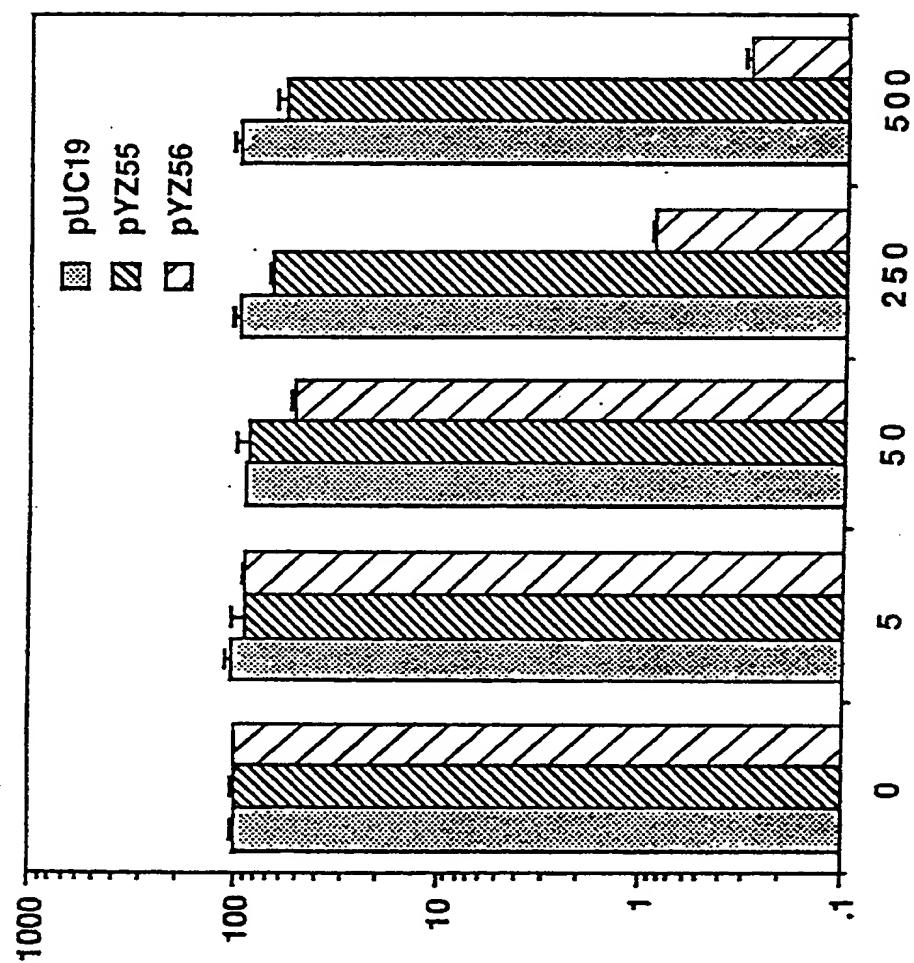
4 / 26

FIG. 2C(2)

## **SUBSTITUTE SHEET**

5 / 26

FIG. 3



SUBSTITUTE SHEET

6 / 26

FIG. 4A

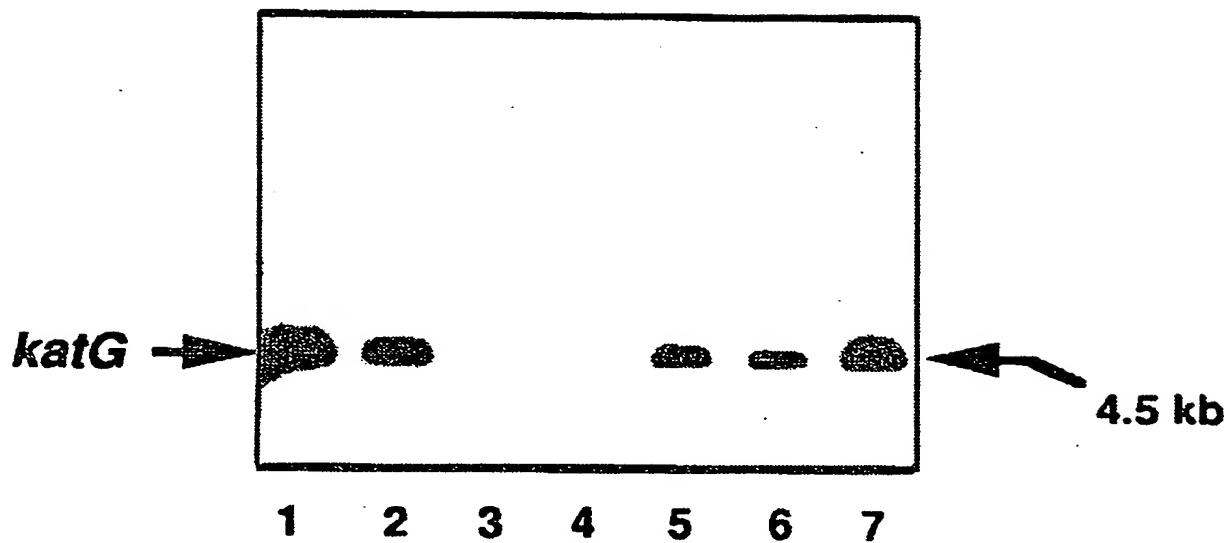
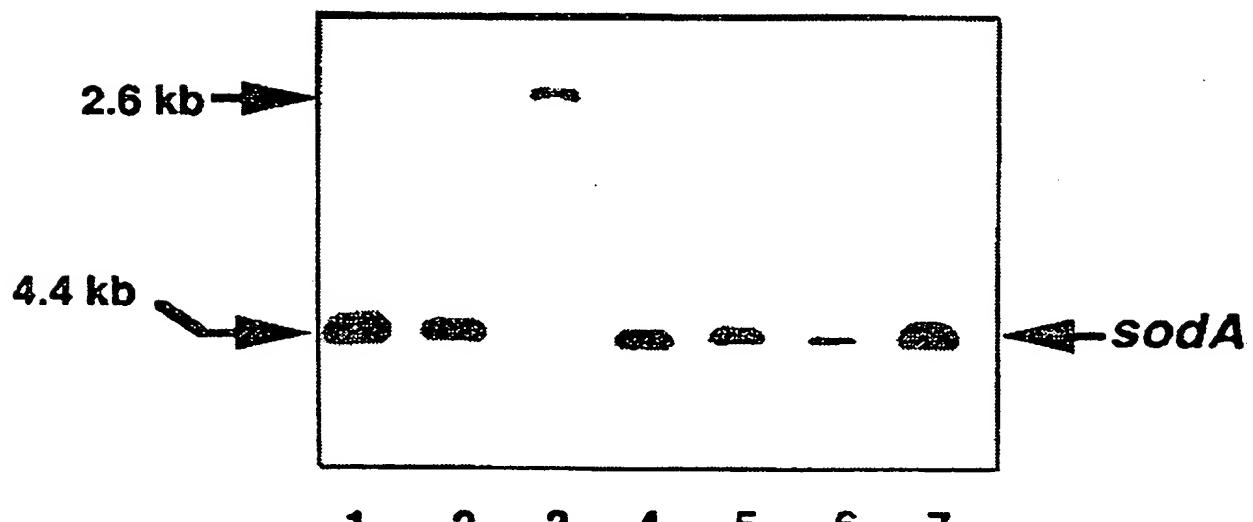


FIG. 4B



SUBSTITUTE SHEET

7 / 26

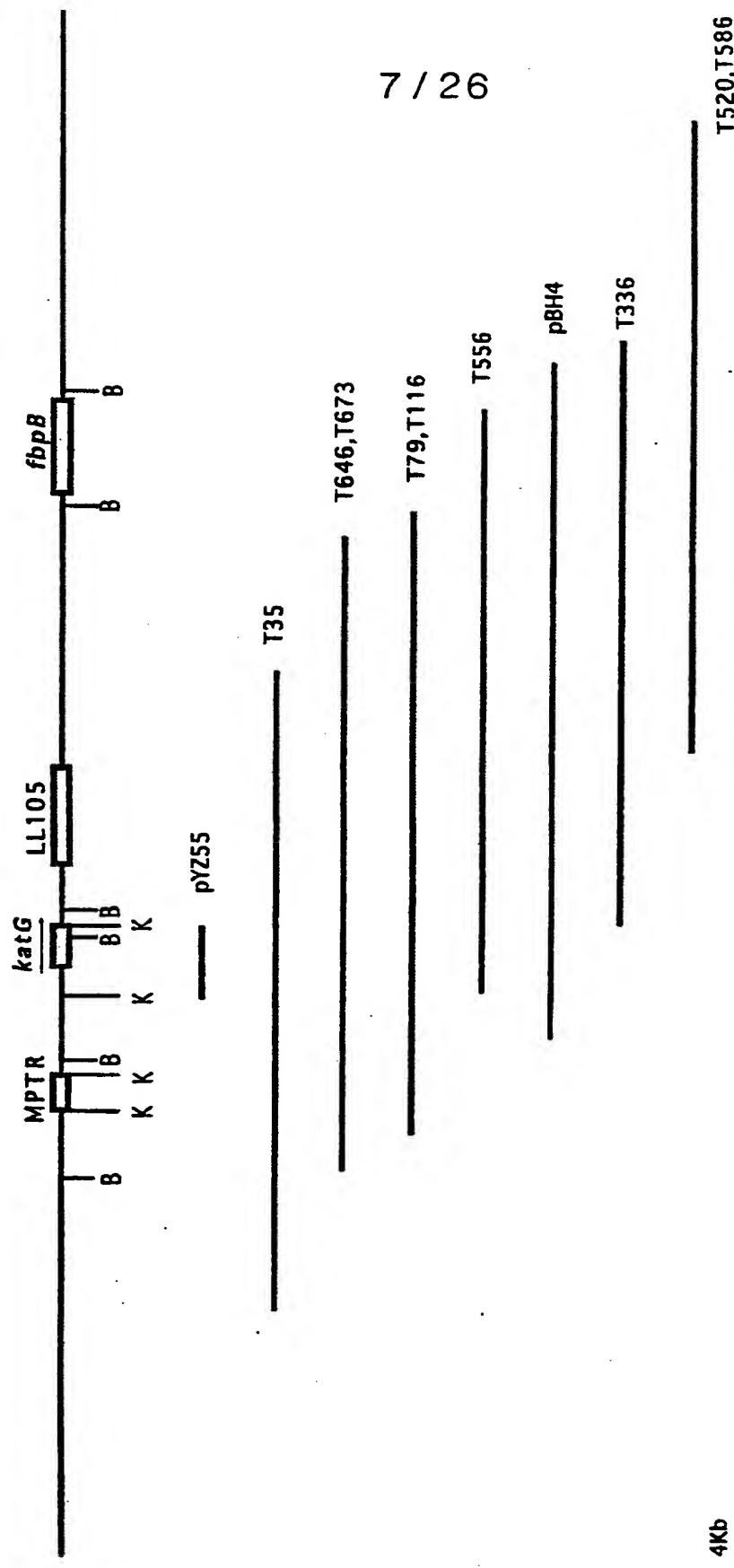


FIG. 5

SUBSTITUTE SHEET

8 / 26

FIGURE 6A(1)

GGTACCGTGA	GGCGATGGT	GGCCGGGGC	CCGGCTGCT	GTAAGGGCG	GGCGAAAC	60
AGCTGTACTC	TGAAATCCCA	TTAGTAACA	ATGTGCTATG	GAATCTCCAA	TGACGAGCAC	120
ACTTCACCGA	ACCCATTAG	CCACCGGGG	GCTGGGCTC	GTACTGGCCG	TGGTGGCTG	180
CGGGGGGG	GGGGGTGACA	GTCGAGAGAC	ACCGCATAAC	GTGCCGAAAG	CGACGACCGT	240
CGACGCAACA	ACGCCGGGC	GGGCCGGGA	GCCACTGACG	ATGCCAGTC	CCATGTTCGC	300
CGACGGGGCC	CCGATCCCGG	TGCAATTAG	CTGCCAAGGGG	GCCAACTGTGG	CCGCCACCGT	360
TGACGTTGTC	GTGCGCCGG	GCGAGCGAAC	TGGCACTCGT	CCTCGATGAC	CCGACGGCGG	420
TCGGGGAACT	GTACGTGCAC	TGGATCGTGA	CCGGAATCGC	CCCTGGCTCT	GGCAGCACGG	480
CGGATGGTCA	GACTCCTGGT	GGTGGGCACA	GGGTGGCCAA	TTCCTGGTGT	GGCAGAAGGAT	540
ACTTCGGTCC	ATGCCCGCC	GCGGGCACCG	GGACACACCA	CTACCGGTTT	ACCTCTTACCC	600
ACCTTCCTGT	CGCGTCCAG	CTGCCACCGG	GAGGCCACGGG	AGTCCAAGGG	GCACAGGGGA	660
TAGCACAGGC	CGCCAGGCAC	AGGCCCGGT	CGTGGCACA	TTCGAAGGGT	GACGCCGGG	720
CATCCCTGGC	GAGGTGGTGTG	AAACCCCTGGC	TTCTCCAAATT	GGGCCCTGGCG	ACAATGATCA	780
ATATGGAATC	GACAGTGGCG	CACGCATTTC	ACCGGATTCG	ACTGGCCATC	TTGGGGCTGG	840
CGCTCCCGT	GGCGCTAGT	GGCTACGGTG	GAAACGGTGA	CAGTCGAAG	GGGGGGCCC	900
TGGGGCGAA	AGCAGCAGCG	CTCGGTGCGA	GTATGCCGA	AACGCCCTAC	GGCGATGTAC	960
TGACAATAG	CAGTCCGGCA	TTCGCCGACG	GTGGGCCGAT	CCCGGAACAG	TACACCTGCA	1020
AAGGAGGCCA	TATCGGGCC	TCCGTTGACC	TGGTGGGCC	CGTTTGGGG	CCCACTCGTT	1080
GTTCGATGATC	CGGACCACTT	CGCGAACCTT	ACGTCCATTG	GATCGTGTATC	GGGATGGCCC	1140
CTGGTGTGG	CAGCAGCCGA	TGGTGGAGCT	CCGGTGGCG	GAATCAGGCT	GGCGAACTCC	1200
AGCGGTCAAGC	CCGCATAACAC	CGGCCCTGTC	CGGCCGGGG	GCACCGGGAC	ACACCAACTAC	1260

SUBSTITUTE SHEET

9 / 26

FIGURE 6A(2)

CGGTTTACCC	TCTTACCACT	TCCTGCCGTG	CCTCCACTCG	CGGACTGGC	TGGCACCAA	1320
GGGGCGGGG	TGATCGGCA	GGCCGCCACC	ATGCCGCC	GGCTCATCGG	AAACATACGAA	1380
GGCTGATCCA	CCCGCCATCC	CACGATCCAG	CGGGCCGGG	CGATGGGTG	CTAGCAGACG	1440
CCTGTCAGC	TAGCCAAAGT	CTTGACTGT	TCCAGAAAAG	GGAGTCATAT	TGTCCTAGTGT	1500
GTCCCTATA	CCGGACTAAG	CCGAACAGCT	CGGAGGGCC	GAACCTGCCG	TGACCCGAC	1560
GGCGGTGCGC	GTCCCTGGAG	CAGTGAATGC	GCATCCACAC	GCGGACACGG	AAACGATT	1620
CGGTCGGT	CGTTTTGCCC	TGCCCCGACGT	ATCCGGCAAG	CCGTGTACGA	CGTGTGCA	1680
GCCCTGACCCG	CCGGGGGCTT	GGTGGAAAG	ATCCAACCC	CGGGCTCCGT	CGGGGCTAC	1740
GAGTCCAGGG	TGGCGACAA	CCACCATCAC	ATCGTCTGCC	GGTCTTGGG	GGTATCGCC	1800
GATGTCGACT	GTGCTGTTGG	CGAGGCACCC	TGTCTGACGG	CCTCGGACCA	TAACGGCTTC	1860
CTGTTGGACG	AGGGGGAGGT	CATCTACTGG	GGTCTATGTC	CTGATTGTC	GATATCCGAC	1920
ACTTCGGAT	CACATCCGTG	ATCACAGCCC	GATAACACCA	ACTCCTGGAA	GGATGCTGT	1980
GCCCAGGCAA	CACCCACCA	TTACAGAAC	CACCACGGG	GCGGCTAGCA	ACGGCTGTCC	2040
CCTCGTGGGT	CATATGAAT	ACCCCGTGA	GGGGGGCGGA	AACCAGGACT	GGTGGCCAA	2100
CCGGCTCAAT	CTGAAAGGTAC	TGCACCAAA	CCGGGGCGTC	GCTGACCCGA	TGGTGGGGC	2160
GTTGACTAT	GCCGGGGAGG	TGGCGACCG	TCGACTTGAC	GCCCTGACCC	GGGACATCGA	2220
GAAAGTGTG	ACACCTCGC	AGCCGTTGGTG	GCCCCCGAC	TACGGCCACT	ACGGGCCGT	2280
GTTTATCCGG	ATGGCGTGGC	ACGGCTGCCGG	CACCTACCGC	ATCCACGACG	GCCGGGGCG	2340
CGCCGGGGC	GGCATGCGAC	GGTTGCCGCC	GCTTAACAGC	TGGCCCGACA	ACGCCAGCT	2400
GGACAAGGGG	CGCCGGCTGC	TGTGGCCGGT	CAAGAAGAAG	TACGGCAAGA	AGCTCTCATG	2460
GGCGGACCTG	ATTGTTTCG	CCGGCAACCG	CTGGCTCGG	AATCGATGGG	CTTCAAGACG	2520
TTCGGGTTCG	GCTTCGGGG	TCGACCAAGT	GGAGACCGAT	GAGGTCTATT	GGGGCAAGGA	2580

SUBSTITUTE SHEET

10 / 26

FIGURE 6A(3)

AGCCACCTGG	CTCGGGATG	ACGGTTACAG	CGTAAGCGAT	CTGGAGAAC	CGCTGGCCGC	2640
GGTGCAGATG	GGCCTGATCT	ACGTGAACCC	GGAGGGCCG	AACGGCAACC	CGGACCCCCAT	27.00
GGCCGGGGCG	GTGACATTC	GCGAGACGTT	TCGGCCATG	GCCATGAACG	ACGTCGAAAC	2760
AGCGGGCTG	ATCGTGGGG	GTCAACATT	CGGTAAGACC	CATGGGCCG	GCCGGCCGA	2820
TCTGGTCGGC	CCCGAACCCG	AGGCTGCTCC	GCTGGAGCG	ATGGCCTGG	GCTGGAAGAG	2880
CTCGTATGGC	ACCGGAACCG	GTAAGGACGC	GATCACCCAGC	GGCATCGAGG	TCTGTATGGAC	2940
GAACACCCCG	ACGAATATGG	ACAAACAGTTT	CCTCGAGATC	CITGTAAGGCT	ACGAGTGGGA	3000
GCTGACCGAAG	AGCCCTGCTG	GCGCTTGGCA	ATACACCGCC	AAGGACGGG	CCGGTGGCCGG	3060
CACCATCCCCG	GACCCGTTCG	GCGGGCCAGG	GGGCTCCCCG	ACGATGCTGG	CCACTGACCT	3120
CTCGCTCGGG	GTGGATCCGA	TCTATGAGCG	GATCACGCGT	CGCTGGCTGG	AACACCCCCGA	3180
GGAAATTGGCC	GACCGAGTCC	GCAAGGCCTG	GTACAAGCTG	ATCCACCGAG	ACATGGGTCC	3240
CGTTCGAGA	TACCTTGGGC	CGCTGGTCCC	CAAGCAGACC	CTGCTGTGGC	AGGATCCGGT	3300
CCCTGGGTC	ACGACGACT	CGTCGGCGAA	GCAGATTGCC	AGCCTTAAGA	GCCAGATCCG	3360
GGCATCGGGA	TIGACTGTCT	CACAGCTAGT	TTGACCGCA	TGGGGCGGG	CGTCGTCGTT	3420
CCGTGGTAGC	GACAAGGGCG	GGGGCCCAA	CGGTGGTCGC	ATCCGCCCTGC	AGCCACAAAGT	3480
CGGTGGGAG	GTCAACGACC	CCGACGGATC	TGGCAAGGT	CATTGGCACCC	CTGAAGAGAT	3540
CCAGGAGTCA	TTCACTCGGC	GGGGAACAT	CAAAGTGTCC	TTCGCCGACC	TCGTCGTGCT	3600
CGGTGGCTGT	GGGCCACTAG	AGAAAGCAGC	AAAGGGGCT	GGCCACAAACA	TCACGGTGCC	3660
CTTCACCCCG	GGCCGGCAAG	ATGCGTCGCA	GGAAACAAACC	GACGTGGAAT	CCTTGTCCGT	3720
GCTGGAGCCC	AAGGGAGATG	GCTTCCGAA	CTACCTCGGA	AAGGGCAACC	GTGCCCCGCC	3780
GAGTACATCG	CTGCTCGACA	AGGGAAACCT	GCTTACGCTC	AGTGGCCCTG	AGATGACGGT	3840
GCTGGTAGGT	GGCCTGGGG	TCCTCGGGCG	AAACTACAAG	CGCTTACCGC	TGGCGTGTGT	3900

SUBSTITUTE SHEET

11 / 26

FIGURE 6A(4)

CACCGAGGCC TCCGAGTCAC TGACCAACGA CTTCTTCGTG AACCTGCTCG ACATGGGTAT 3960  
CACCTGGAG CCCTGCCAG CAGATGACGG GACCTACCA GGCAGGTG GCAGTGGCAA 4020  
GGTGAAGTGG ACCGGCAGCC GCGTGGACCT GGTCTTCGGG TCCAACTCGG AGTTCGGGC 4080  
GCTTGTGAG GTCTATGCC CGATGACGG GCAGGCAAG TTCGTGACAG GATTCGTGCC 4140  
TGCCTGGAC AAGGTGATGA ACCTCGACAG GTTCGACGTG CGCTGATTG GGTGATCGG 4200  
CCCTGCCGC CGATCAACCA CAACCCGCC CAGCACCCCG CAGCACCCCG CGAGCTGACCC GGCTCGGGG 4260  
GTTGCTGGTGT TTGCCCCGG CGATTTGTCA GACCCCCGT GCATGGTGT CGCACGGACG 4320  
CACGAGACGG GGATGACGAG ACGGGGATGA GGAGAAAGGG CGCCGAAATG TGCTGGATGT 4380  
GGGATCACCC GGAAGCCACC GCGGAGGAGT ACCTCGACGA GGTGTACGGG ATATGCTCA 4440  
TGCATGGCTG GCGGTACAG CACGGGAGT GCGAGGGACG GCCATTGGC TACACGGTTG 4500  
GTCCTAACCCG GCGGGCTTG CCCGAACTGG TGGTGAATGG CCTCTCGCCA CGACGTGGGC 4560  
ACGGGTTGTT GAAACATGCC CGAAGGGCTC TGGTGGTGA CTTGCTGACT CCCGGTATGT 4620  
AGACCACCT CAAAGCCGGC CCTCTTGTCTG AAACGGTCCA GGCTACACAT CGGACGGGC 4680  
ATTTGTATTG TGGGATGCC ATCTTTGGC ACAAGGGTAC GGCCTTGCAG TTGGTGGGG 4740  
CGGACCGGGT GGTGCTGGC CGTGGGGGG GGACTTCGAC GAAGGTGGC GTACCC 4795

SUBSTITUTE SHEET

12 / 26

FIGURE 6B(1)

## **SUBSTITUTE SHEET**

13 / 26

FIGURE 6B(2)

1241	1251	1261	1271	1281	1291
CAGCCTGCCG	AACTCCAGCG	GTCAAGCCCG	ATACACCGGC	CCCTGCCCGC	CGGGGGCAC
***	***	***	***	***	***
CAGCGTGCCG	AATTCTGGTG	GTCGGGCAAGG	ATACTTCGGT	CCATGCCGC	CGGGGGCAC
541	551	561	571	581	591
1301	1311	1321	1331	1341	1351
CGGGACACAC	CACTACCGGT	TTACCCCTCTA	CCACCTTCCT	GGCCTGCCCT	CA-CTTCGGC--
***	***	***	***	***	***
CGGGACACAC	CACTACCGGT	TTACCCCTCTA	CCACCTTCCT	GTGGGCC-TC	CAGCT-GCCA
601	611	621	631	641	651
1361	1371	1381	1391	1401	1411
--GGGACTGG	CT--GGGA--	CACAAGGGGC	GGGGGTGATC	GGGCAGGGCCG	CCACCATG-C
***	*	***	***	***	*
CGGGGA--G	CCACGGGAGT	C-CAGGGGC	ACAGGGATA	GCACAGGGCCG	CCAGC--GAC
661	671	681	691	701	711
1421	1431	1441	1451	1461	-----
AGGCCCGGCT	CATCGGAACA	TACGAAGGCT	GATCCACCCG	CCATCCC	-----
721	731	741	751	761	-----

**SUBSTITUTE SHEET**

14 / 26

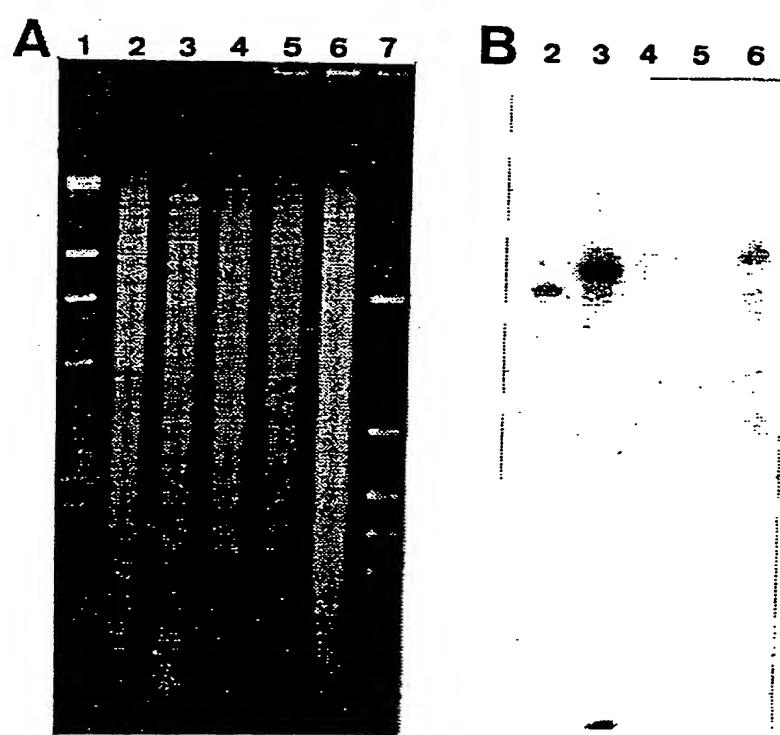


FIG. 7

**SUBSTITUTE SHEET**

1 MTKATG MPEQHPPITE TTTGAASNGC PVIYGHMKYPV EGGGNODWWP NPAVADPMGA AFDYAAEVAT  
 ECKATG MSTSSDDIH NTTATGKCPF HQGGHDQSAQ AGTTTRDWWP HSNRSNPILGE DFDYRKEF  
 STKATG MSTTDDTH NTLSTGKCPF HQGGHDRSAG AGTASRDWWP HSNRSNPILGE DFDYRKEF  
 BSPERA MENQ NRQNAAQCPF HEVTNQSS. NRTTNKDWMP NQNLNLILHQ HDRKTNPHDE ENYAAEFQ.  
 CCP CONSENSUS --MSTI-DDTH NTT---KCPF HQGGHDQSAQ AGTTTRDWWP NQI--DLLHQ HSNRSNPLGE DFDY-KEF--

70

71 MTKATG SRLD.:::ALT RDIEEVMTS QFWWPADYGH YGPLFIRMAW HAAGTYRIHD GRGGAGGGM Q RFAPLNSWPD  
 ECKATG SKLDYY.GLK KDLKALLTES QFWWPADWGS YAGLFLIRMAW HGAGTYRSID GRGGAGRGQQ RFAPLNSWPD  
 STKATG SKLDYY.SALK GDLKALLTDS QFWWPADWGS YVGLFLIRMAW HGAGTYRSID GRGGAGRGQQ RFAPLNSWPD  
 BSPERA KLDYW.ALK EDLRKLMTES QDWWPADYGH YGPLFIRMAW HSAGTYRIGD GRGGASTGTO RFAPLNSWPD  
 CCP KVNAIAALK .KVNIAALK .LRED DEX. DNY. YGPLVYRLAW HISGTMWDKHD MTGGSYGGTY RFKKEFNDPS  
 CONSENSUS SKLDYY-ALK .-DLKALLTES QFWWPADYGH YGPLFIRMAW HGAGTYR--D GRGGAG-G-Q RFAPLNSWPD

140

141 MTKATG NASLDKARRL LWPVKKKYGK KISWADLIVF AGNRCARNRW ASRRSGSASG .::: VDOWETD .: EYVGKEAT  
 ECKATG NVSLDKARRL LWPPIKOKYGO KISWADLFL AGNVALENSG FRTFGFGAGR .::: .EDYWEPD LDVNWGDEKA  
 STKATG TYSLDKARRL LWPPIKOKYGO KISWADLFL AGNVALENSG FRTFGFGAGR .::: .EDYWEPD LDVNWGDEKA  
 BSPERA NANLDKARRC YGRSKRNTGT K.SLGPICSF WRAMSLLNRW VEKRILDAAG FLTSGIRKKT FIGDRKKSGS  
 CCP NAGLQNGFKF LEPIKKEFP. WISSGDLFLS LGGTAVOEMQ GPKIPWRCGR VDTPEDTPD  
 CONSENSUS NASLDKARRL LWPPIK-KYGO KISWADLFL AGNVALEN-- FR--GF-AGR --TEDVWEPD LDVNWG-EKA

210

211 MTKATG WLGDDGYSVS DLENPLAAVO MGLLIVNPEA PNGNPDPMAA AVDIRETFRR MAMNDVETAA LIVGGHTFGK  
 ECKATG WLTHR.HPEA LAKAPLGLATE MGLLIVNPEG PDHSGEPLSA AAAIRATFGN MGMNDEETVA LIAGGHTLGK  
 STKATG WLTHR.HPEA LAKAPLGLATE MDLLIVNPEG PNHSGEPLSA AAAIRATFGN MGMNDEETVA LIAGGHTLGK  
 BSPERA PLNAPVIAS SKTRSPERANG VNLRQPRRAG RQAGSKSRGT SA.::: ETFRR MGMNDEETVA LIAGGHTLGK  
 CCP RL. LEPIKKEFP. WISSGDLFLS AGYVRTFFQRLNMMNDREVVA LM.GAHALGK  
 CONSENSUS WLTHR-HPE- LAKAPLGLATE MGLLIVNPEG PNHS--PLSA AAAIR-TF-R MGMNDEETVA LIAGGHTLGK

281

282 MTKATG THGAGPADLV GPEPEAAPPLE QMGLGWKSSY GTGTGKDAIT SCIEVVNTNT PTKWDNSFILE ILYGEWELL  
 ECKATG THGAGPTSNV GPDPEAAPPLE EQLGWASTY GSGVGADAIT SGLEVVWTOT PTQWSNYFFE NLFKYEWVQT  
 STKATG THGPAASHV GADPEAAPPLE AQGLGWASSY GSGVGADAIT SGLEVVWTOT PTQWSNYFFE NLFKYEWVQT  
 BSPERA AHRGGPATHV GPEPEAAPPLE AQGLGWISSY GKGKGSDTIT SGIEAWTPT PTQWDTSYFD MLEFGDWWLT  
 CCP TH.GG.PASHV GP-PEAAPPLE AQGLGWASSY GSGVGADAIT SGYEGPWGAA NNVFTNEFYL NLLNEDWKL  
 CONSENSUS TH(275) SG-EVVWTOT PTQW-N-FFE NLF-YEWVLT

350

351 MTKATG KSPAGAWOYT AKDGAGAGTI PDPFGGPGR .::: SPTMLATD LSLRVDPIYE RITRRWLEHP EELADEFKA  
 ECKATG RSPAGAIQFE AVEII PDPFDPSKK .::: KPTMVLTD LTLLRFDPFE KISRRFLNDP QAFNEAFARA  
 STKATG RSPAGAIQFE AVE. APDI PDPFDPSKK .::: KPTMVLTD LTLLRFDPFE KISRRFLNDP QAFNEAFARA  
 BSPERA KSPAGAWOYT AVDPDEKDILA PDAEDPSKK .::: VPTMMTTD LALRFDPFEYE KIARRFHONP EEEFAAFARA  
 CCP KNDANNEQND SKSGY .::: MLLPTD YSLIQDPKYL SIVKEYANDQ DKEFKDFSKA  
 CONSENSUS KSPAGA-Q-E AVDG-APDI PDPFDPSKK .::: KPTMVLTD L-IRFDPEYE KISRRFLNDP E-F-EAFARA

420

D(380)

## FIGURE 8(2)

421	WYKLTHRDW.	GPVARYL GPIUPKOTILL	WODPVPAVST	ISSAKOIASL	KSOIRASGLT	VSOLVUSTAWA	490
{MTKATG	WFKLTHRDW.	... GPKSRYI	GP EVPKEDLI	WODPLPOPIY	NPTEODIDL	KFAIAASGLS	VS E L V S V A W A
{ECKATG	WFKLTHRDW.	... GPKSRYI	GP EVPKEDLI	WODPLPOPIY	OPTOEDIIL	KAAIAASGLS	VS E M V S V A W A
{STKATG	WFKLTHRDW.	... GPKSRYI	GP EVPKEDLI	WODPLPOPIY	ELTEAEIEI	KAKILNSGLT	VS E L V K T A W A
{BSPERA	WFKLTHRDW.	... GPKSRYI	GP EVPKEDLI	WODPLPOPIY	WQDPIPEVDY		
{CCP	FEKLENGIT	FPKDAPSFI	FPKLEEGGL.	WQDP-PQ--Y	-PTE-DII-L	KAAIAASGL-	VS E L V S - A W A
CONSENSUS	WFKLTHRDW-	---GPK-RYI	GP EVPKEDLI	WQDP-PQ--Y	-PTE-DII-L	KAAIAASGL-	VS E L V S - A W A
491	AASSFRGSDK	RGGI	RLOPOVGWEV	NDPDGSAOGH	SHPEEIQESF	TRRGNIKVSF	560
{MTKATG	SASTFRGDK	NGGRI	ALMPORDWDV	N...AAAVRL	PVLEIQL	ADLVVIGGCA	
{ECKATG	SASTFRGDK	NGGRI	ALAPORDWDV	N...AAAVRL	PVLEIQL	STTESSLIDK	
{STKATG	SASTFRGDK	NGGRI	ALAPORDWDV	N...AAAVRL	PVLEIQL	ADLIVLAGV	
{BSPERA	SAA.	RSATR	ISAATNGRRI	NEPERLAKVL	SVLRGHPA...	...KTTNAKS	
{CCP	SASTFRGDK	RGGI-NGAR-	-LAPORDW-V	N-P--AARVL	VLLEEIQ---	...R TAEKS	
CONSENSUS	SASTFRGDK	RGGI-NGAR-	-LAPORDW-V	N-P--AARVL	VLLEEIQ---	...R TAEKS	
561	PLEKAAKAG	HNTIVPF...	TPGPHDASQE	OTDVEFAVL	EPKADGFRN.	... YLGKGNR	630
{MTKATG	GVEKAASAAG	LSIHVPF...	APGRVDAEOD	OTDIEMFELL	EPIADGFRN.	... YRARL DV	
{ECKATG	GIEQAAAAR	LSIHVPF...	PPGRVDAEOD	OTDIEMFELL	EPIADGFRN.	STTESSLIDK	
{STKATG	VSIHVPF...	LSIHVPF...	PPGRVDAEOD	OTDIEMFELL	EPIADGFRN.	ADLIVLAGV	
{BSPERA	RWKRPATPA	LSMKCHFSIA	AMRHKS KSKPM	SKALPCWNRQ	QMASATIKSK	...KTTNAKS	
{CCP	G-EKAAAAG	LSIHVPF--	APGR-DARQD	OTDIEMF-LL	EP IADGFRN-	...R TAEKS	
CONSENSUS	G-EKAAAAG	LSIHVPF--	APGR-DARQD	OTDIEMF-LL	EP IADGFRN-	...R TAEKS	
631	ANLTLTSAPE	MTVLVGGGLRV	LGANYKRLPL	GVFTEASESL	TNDFFVNLLD	MGITWEPSA	700
{MTKATG	AQQLLTAPE	MTLVGGMVR	LGGNFDGSKN	GVFTDRGVVL	SDFFVNLLD	DDGTYQCKD	
{ECKATG	AQQLLTAPE	MTLVGGMVR	LGTNFDGQN	GVFTDKPGVL	SDFFVNLLD	MRYEWKATDE	
{STKATG	AQQLLTAPE	MTLVGGMVR	LGTNFDGQN	GVFTDKPGVL	SDFFVNLLD	SKELFEGRDR	
{BSPERA	PRNDGLSWR.	... . . . . FAR	VGPNRYRHLBH	GVFTDRGVVL	SDFFVNLLD	ANELFEGRDR	
{CCP	AQQLLT-APE	MTVLVGGMVR	LG-N-DG-PN	GVFTDR-GVL	-NDFFVNLLD	MRYEWKPTDS	
CONSENSUS	AQQLLT-APE	MTVLVGGMVR	LG-N-DG-PN	GVFTDR-GVL	-NDFFVNLLD	.. . . . . GIVEIRDR	
701	GSGKVKWTGS	RVDLFGSNS	ELRALVEVYA	PMTRQAKFVT	MGITWEPSA	DDGTYQCKD	767
{MTKATG	ETGEVKFTAS	RADLFGSNS	VIRAVAEVYA	SSDAHEKFVK	NDFFVNLLD	MRYEWKATDE	
{ECKATG	LTGEVKFTAS	RADLFGSNS	VIRALAEVYA	CSDAHEKFVK	SDFFVNLLD	SKELFEGRDR	
{STKATG	LTGEVKFTAS	RADLFGSNS	VIRALAEVYA	QDDNQEKFVR	SDFFVNLLD	ANELFEGRDR	
{BSPERA	KTGEVRWTAT	RVDLFGSNS	VIRSYAEFYA	DFINAVVKVM	SDFFVNLLD	MRYEWKPTDS	
{CCP				NADRFDLVKK	SDFFVNLLD	.. . . . . GIVEIRDR	
CONSENSUS	-TGEVKFTA-	R-DLVFGSNS	VIRALAEVYA	-SDA-EKFVK	SDFFVNLLD	.. . . . . GIVEIRDR	

17 / 26

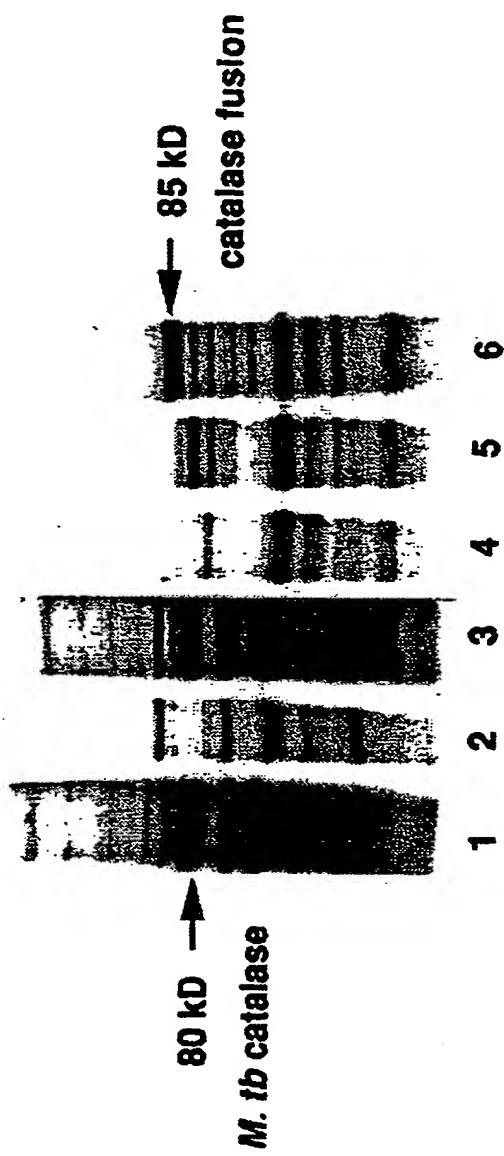


FIG. 9

SUBSTITUTE SHEET

18 / 26

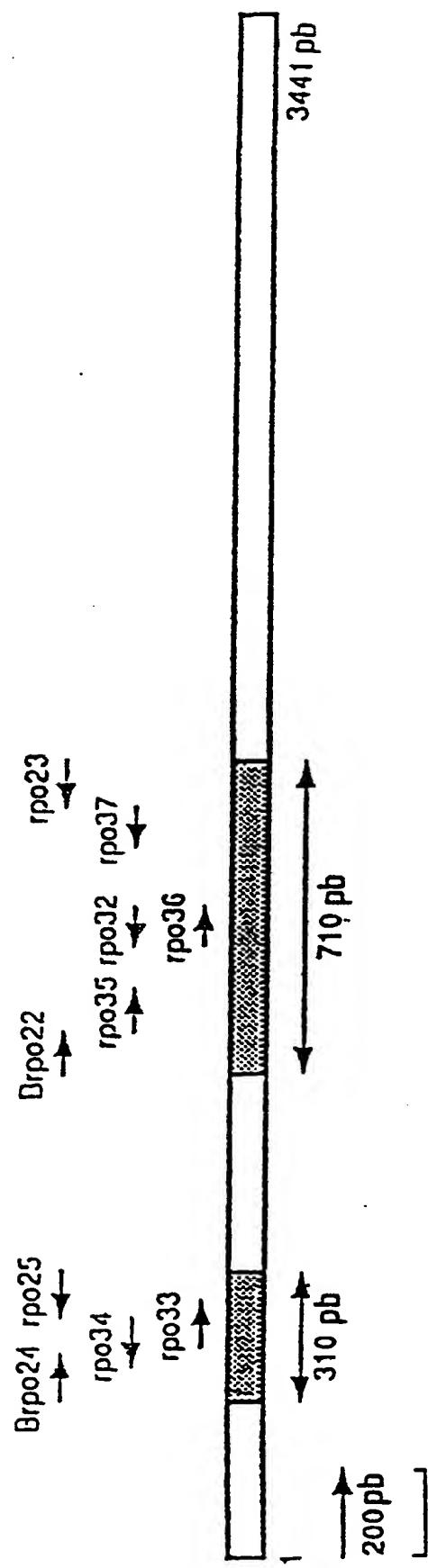


FIG. 10

SUBSTITUTE SHEET

19 / 26

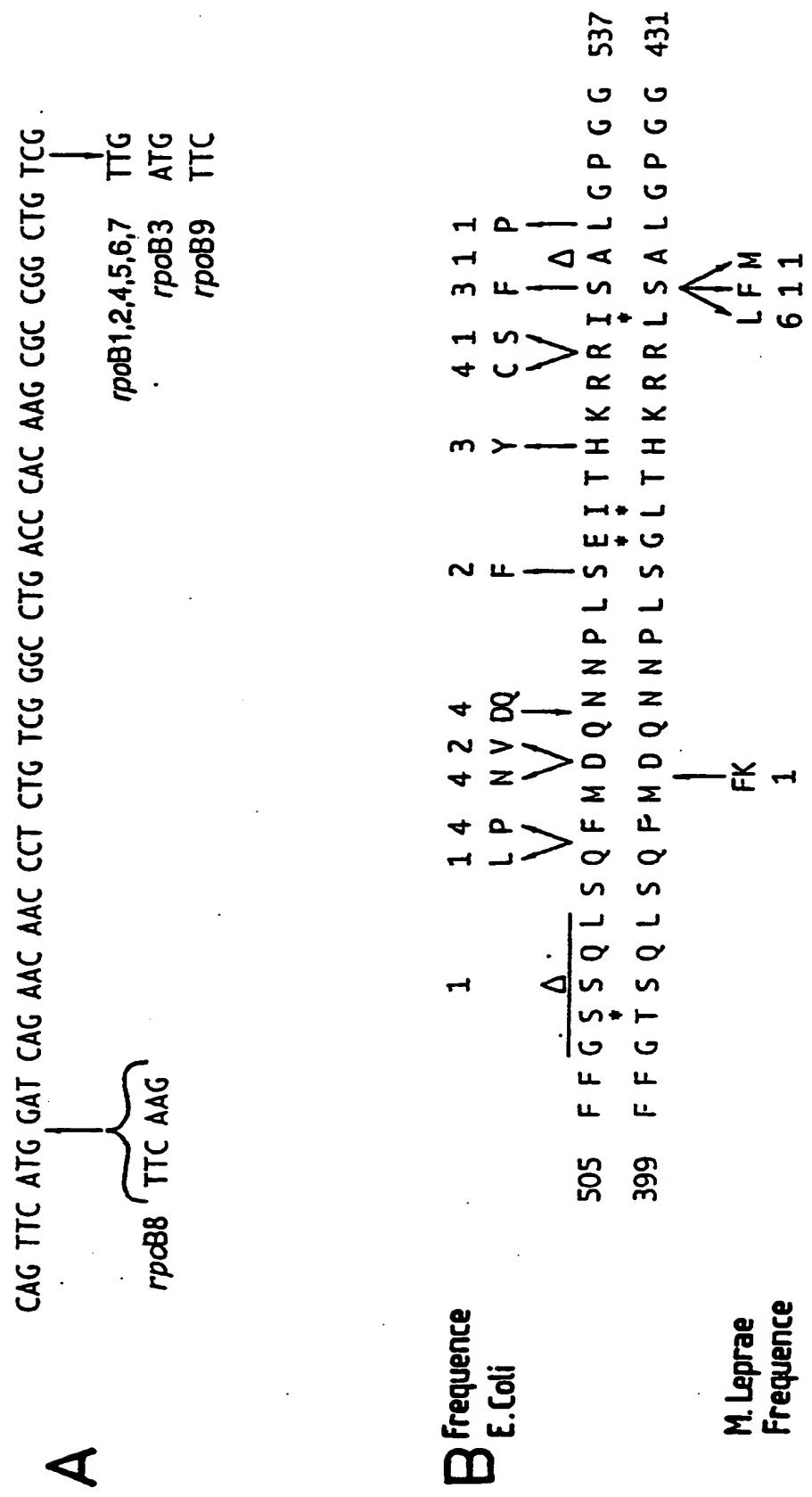


FIG. 11

SUBSTITUTE SHEET

20 / 26

MetProGlyAlaProAsnArgIleSerPheAlaLysLeuArgGluProLeuGluValPro	60
GTGCCCGCGGCCAACCGAACCGCTTGAGGTTCCG	
GlyLeuLeuAspValGlnThrAspSerPheGluTrpLeuIleGlySerProCysTrpArg	120
GGGCTACTTGATGTGCAGACTGATTGAGTGGTTGATCGGATGCCGTGCTGGCGT	
AlaAlaAlaAlaAlaSerArgGlyAspLeuLysProValGlyGlyLeuGluGluValLeuTyr	180
GCAGCGGCCGCAAGCCGGCGATCTCAAGCCGGTGGTGGTCTCGAAGAGGTGCTCTAC	
GluLeuSerProIleGluAspPheSerGlySerMetSerLeuSerPheSerAspProArg	240
GAGCTGTCGCCGATCGAGGATTCTCCGGCTCAATGTCATTGTCATTCTCCGATCCCCGT	
PheAspGluValLysAlaProValGluGluCysLysAspLysAspMetThrTyrAlaAla	300
TTTGACGAAGTCAAGGCGCCCGTCGAAGAGTGCAGAACAGACATGACGTACGCGGCC	
ProLeuPheValThrAlaGluPheIleAsnAsnAsnThrGlyGluIleLysSerGlnThr	360
CCGCTGTTCGTCACGGCCGAGTTCATCAACAAACACCGGGGAGATCAAGAGCCAGACG	
ValPheMetGlyAspPheProMetMetThrGluLysGlyThrPheIleIleAsnGlyThr	420
GTGTTATGGCGACTTCCCTATGACTGAGAAGGAAACCTTCATCATCACACGGGACC	
GluArgValValValSerGlnLeuValArgSerProGlyValTyrPheAspGluThrIle	480
GAGCGTGTCTCGTTAGCCAGCTGGTGCCTCCCTGGAGTATACTCGACGAGACGATC	
AspLysSerThrGluLysThrLeuHisSerValLysValIleProSerArgGlyAlaTrp	540
GACAAGTCCACAGAAAAGACGCTGCATAGTGTCAAGGTATTCCAGCCGGTGCCTGG	
LeuGluPheAspValAspLysArgAspThrValGlyValArgIleAspArgLysArgArg	600
TTGGAATTCGATGTCGATAAACCGCGACACCGTCGGTGTCCGCATTGACCGGAAGCGCCGG	
GlnProValThrValLeuLeuLysAlaLeuGlyTrpThrSerGluGlnIleThrGluArg	660
CAACCCGTCACGGTCTTCTCAAAGCGCTAGGTTGGACCAGTGAGCAGATACCGAGCGT	
PheGlyPheSerGluIleMetArgSerThrLeuGluLysAspAsnThrValGlyThrAsp	720
TTCGGTTTCTCCGAGATCATGCGCTCGACGCTGGAGAACACACAGTTGGCACCGAC	
GluAlaLeuLeuAspIleTyrArgLysLeuArgProGlyGluProProThrLysGluSer	780
GAGGCGCTGCTAGACATCTATCGTAAGTTGCGCCCAGGTGAGCCGCCACTAAGGAGTCC	
AlaGlnThrLeuLeuGluAsnLeuPhePheLysGluLysArgTyrAspLeuAlaArgVal	840
GCGCAGACGCTGTTGGAGAACCTGTTCTCAAGGAGAACGCTACGACCTGGCCAGGGTT	
GlyArgTyrLysValAsnLysLysLeuGlyLeuHisAlaGlyGluLeuIleThrSerSer	900
GGTCGTTACAAGGTCAACAAAGCTCGGGTGCACGCCGGTGAGTTGATCACGTCGTCC	
ThrLeuThrGluGluAspValValAlaThrIleGluTyrLeuValArgLeuHisGluGly	960
ACGCTGACCGAAGAGGATGTCGTCGCCACCATAGAGTACCTGGTTCGTCTGCATGAGGGT	
GlnSerThrMetThrValProGlyGlyValGluValProValGluThrAspAspIleAsp	1020
CAGTCGACAATGACTGTCCCAGGTGGGGTAGAAGTGCCAGTGGAAACTGACGATATCGAC	
HisPheGlyAsnArgArgLeuArgThrValGlyGluLeuIleGlnAsnGlnIleArgVal	1080
CACTTCGGCAACCGCCGGCTGCGCACGGTCGGCAATTGATCCAGAACCGAGATCCGGGTC	
GlyMetSerArgMetGluArgValValArgGluArgMetThrThrGlnAspValGluAla	1140
GGTATGTCGCGGATGGAGCGGGTGGTCCGGGAGCGGATGACCACCCAGGACGTCGAGGCG	

FIG. 312 (1)

21 / 26

IleThrProGlnThrLeuIleAsnIleArgProValValAlaAlaIleLysGluPhePhe ATCACGCCGCAGACGCTGATCAATATCCGTCCGGTGGTCGCCGCTATCAAGGAATTCTTC	1200
GlyThrSerGlnLeuSerGlnPheMetAspGlnAsnAsnProLeuSerGlyLeuThrHis GGCACCCAGCCAGCTGTCGAGTCAGTTCATGGATCAGAACAAACCTCTGTCGGGCCTGACCCAC	1260
LysArgArgLeuSerAlaLeuGlyProGlyGlyLeuSerArgGluArgAlaGlyLeuGlu AAGCGCCGGCTGTCGGCGCTGGGCCCGGTGGTTGTCGGTGTGAGCGTGCCGGCTAGAG	1320
ValArgAspValHisProSerHisTyrGlyArgMetCysProIleGluThrProGluGly GTCCGTGACGTGCACCCTCGCACTACGGCCGGATGTGCCGATCGAGACTCCGGAGGGC	1380
ProAsnIleGlyLeuIleGlySerLeuSerValTyrAlaArgValAsnProPheGlyPhe CCGAACATAGGTCTGATCGGTTCATGTCGGTGTACGCGCGGGTCAACCCCTCGGGTTTC	1440
IleGluThrProTyrArgLysValValAspGlyValValSerAspGluIleGluTyrLeu ATCGAAACACCGTACCGCAAAGTGGTTGACGGTGTGGTCAGCGACGAGATCGAAATACTTG	1500
ThrAlaAspGluGluAspArgHisValValAlaGlnAlaAsnSerProIleAspGluAla ACCGCTGACGAGGAAGACCGCCATGTCGTGGCGCAGGCCACTCGCCGATCGACGAGGCC	1560
GlyArgSerSerSerArgAlaCysTrpValArgArgLysAlaGlyGluValGluTyrVal GGCCGTTCCCTCGAGCCGCGTGTGGTGCAGCCAGGCGAGGTGGAGTACGTG	1620
AlaSerSerGluValAspTyrMetAspValSerProArgGlnMetValSerValAlaThr GCCTCGTCCGAGGTGGATTACATGGATGTCGCCACGCCAGATGGTGTGGTGGCCACA	1680
AlaMetIleProPheLeuGluHisAspAspAlaAsnArgAlaLeuMetGlyAlaAsnMet GCGATGATTCCGTTCCCTTGAGCACGACGCCAACCGTGGCTGATGGGATGcgCTAACATG	1740
GlnArgGlnAlaValProLeuValArgSerGluArgProLeuValGlyThrGlyMetGlu CAGcgCCAAGCGGTTCCGTTGGTGCAGCGAACGACCGTTGGTGGGTACCGGTATGGAG	1800
LeuArgAlaAlaIleAspAlaGlyHisValValValAlaGluLysSerGlyValIleGlu TTGCGCGCGGCCATCGACGCTGCCACGTCGTTGGAGAAGTCCGGGTGATCGAG	1860
GluValSerAlaAspTyrIleThrValMetAlaAspAspGlyThrArgArgThrTyrArg GAGGTTTCCGCCACTACATCACCCTGATGGCCGATGACGGCACCCGGGACTTATCGG	1920
MetArgLysPheAlaArgSerAsnHisGlyThrCysAlaAsnGlnSerProIleValAsp ATGCGTAAGTTCGCGCGCTCCAACCACGGCACCTGCCAACAGTCCCGATCGTGGAT	1980
AlaGlyAspArgValGluAlaGlyGlnValIleAlaAspGlyProCysThrGluAsnGly GCGGGGGATCGGGTCGAGGCCGGCCAAGTGATTGCTGACGGTCCGTGCACTGAGAACGGC	2040
GluMetAlaLeuGlyLysAsnLeuLeuValAlaIleAsnAlaValGlyGlySerThrThr GAGATGGCGTTGGCAAGAACTTGCTGGTGGCGATCAATGCCGTGGAGGGTCAACAACT	2100
AsnGluAspAlaIleIleLeuSerAsnArgLeuValGluGluAspValLeuThrSerIle AACGAGGATGCGATCATCCCTGTAACCGACTGGTCGAAGAGGGACGTGCTTACTTCGATT	2160
HisIleGluGluHisGluIleAspAlaArgAspThrLysLeuGlyAlaGluGluIleThr CACATTGAGGAGCATGAGATCGACGCCGTGACACCAAGCTGGTGCTGAGGAGATCACC	2220
ArgAspIleProAsnValSerAspGluValLeuAlaAspLeuAspGluArgGlyIleVal CGGGACATTCCCAACGTCTCCGATGAGGTGCTAGCCGACTTGGACGAGCGGGCATCGTG	2280
ArgIleGlyAlaGluValArgAspGlyAspIleLeuValGlyLysValThrProLysGly CGGATTGGCGCGGGAGGTTGACGGTGATATCCTGGTTGGCAAGGTACCCCGAAGGGG	2340

FIG. 3 suite 12(2)

GluThrGluLeuThrProGluGluArgLeuLeuArgAlaIlePheGlyGluLysAlaArg GAAACTGAGCTGACACCGGAAGAGCGGTTGCTGCGGGCGATCTCGGCAGAAAGGCCGC	2400
GluValArgAspThrSerLeuLysValProHisGlyGluSerGlyLysValIleGlyIle GAGGTCCGTGACACGTCGCTGAAGGTGCCACACGGCGAATCCGGCAAGGTGATCGGCATT	2460
ArgValPheSerHisGluAspAspAspGluLeuProAlaGlyValAsnGluLeuValArg CGGGTGTCTCCCATGAGGATGACGACGAGCTGCCCGCCGGTCAACGAGCTGGTCCGT	2520
ValTyrValAlaGlnLysArgLysIleSerAspGlyAspLysLeuAlaGlyArgHisGly GTCTACGTAGCCCAGAAGCGCAAGATCTCTGACGGTGACAAGCTGGCTGGCGGACGGC	2580
AsnLysGlyValIleGlyLysIleLeuProAlaGluAspMetProPheLeuProAspGly AACAAAGGGCGTGATCGGCAAGATCCTGCCTGCCGAGGATATGCCGTTCTGCCAGACGGC	2640
ThrProValAspIleIleLeuAsnThrHisGlyValProArgArgMetAsnValGlyGln ACCCCGGTGGACATCATCCTAACACTCACGGGGTGCCGGCGGATGAACGTCGGTCAG	2700
IleLeuGluThrHisLeuGlyTrpValAlaLysSerGlyTrpLysIleAspValAlaGly ATCTGGAAACCCACCTGGGTGGTAGCCAAGTCCGGCTGGAAAGATCGACGTGGCCGGC	2760
GlyIleProAspTrpAlaValAsnLeuProGluGluLeuHisAlaAlaProAsnGln GGTATACCGGATTGGCGGTCAACTTGCGCTGAGGAGTTGTTGCACGCTGCGCCCCAACAG	2820
IleValSerThrProValPheAspGlyAlaLysGluGluLeuGlnGlyLeuLeuSer ATCGTGTGACCCCCGGTGTGACGGCGCCAAGGAAGAGGAACACTACAGGGCCTGTTGTC	2880
SerThrLeuProAsnArgAspGlyAspValMetValGlyGlyAspGlyLysAlaValLeu TCCACGTTGCCAACCGCGACGGCGATGTGATGGTGGCGCGACGGCAAGGGCGGTGCTC	2940
PheAspGlyArgSerGlyGluProPheProTyrProValThrValGlyTyrMetTyrIle TTCGATGGGCGCAGCGGTGAGCCCTATCCGGTACGGTTGGCTACATGTACATC	3000
MetLysLeuHisHisLeuValAspAspLysIleHisAlaArgSerThrGlyProTyrSer ATGAAAGCTGCACCACTTGGTGGACGACAAGATCCACGCCGCTCCACCGGCCGTACTCG	3060
MetIleThrGlnGlnProLeuGlyGlyLysAlaGlnPheGlyGlyGlnArgPheGlyGlu ATGATTACCCAGCAGCCGGTGGTGTAAGGCACAGTTGGTGGCCAGCGATTGGTGAG	3120
MetGluCysTrpAlaMetGlnAlaTyrGlyAlaAlaTyrThrLeuGlnGluLeuLeuThr ATGGAGTGCTGGGCCATGCAGGCCTACGGTGCGGCCACAGCTGCAGGAGCTGTTGACC	3180
IleLysSerAspAspThrValGlyArgValLysValTyrGluAlaIleValLysGlyGlu ATCAAGTCCGACGACACCGTCGGTCAAGGTTACGAGGCTATCGTTAAGGGTGAG	3240
AsnIleProGluProGlyIleProGluSerPheLysValLeuLeuLysGluLeuGlnSer AACATCCCCGAGCCGGCATCCCCGAGTCGTTCAAGGTGCTCAAGGAGTTACAGTCG	3300
LeuCysLeuAsnValGluValLeuSerSerAspGlyAlaAlaIleGluLeuArgGluGly CTGTGTCTAACGTCGAGGTGCTGTCGTCGGTACGGTGCAGGCGATCGAGTTGCGCGAAGGT	3360
GluAspGluAspLeuGluArgAlaAlaAsnLeuGlyIleAsnLeuSerArgAsnGlu GAGGATGAGGACCTCGAGCGGGCTGCGGCCAACCTCGGTATCAACTTGCCCGAACGAA	3420
SerAlaSerIleGluAspLeuAla*** TCGGCGTCCATAGAAGATCTGGCTTAG	3447

FIG.3 suite 12(3)

60 GlyAsnArgLeuArgThrValGlyGluLeuIleGlnAsnGlnIleArgValGlyMet  
GGAACCGGCCCTGCGTACGGTGGCTGAGCTGATCCAAACAGATCCGGTGGCATG  
120 SerArgMetGluArgValValArgGluArgMetThrThrGlnAspValGluAlaIleThr  
TCGGGGATGGAGCCGGTGGCTGGAGGGATGACCAACCCAGGACGTGGAGGGATCACA  
180 ProGlnThrLeuIleAsnIleArgProValValAlaAlaIleLysGluPhePheGlyThr  
CCGAGACGTTGATCAACATCCGGCCGGTGGCTGGCGATCAAGGAGTTCTTCCGGCAC  
240 SerGlnLeuSerGlnPheMetAspGlnAsnProLeuSerGlyLeuThrHisLysArg  
AGCCAGCTGAGCCAATTATGGACCAACAAACCCGGTGGCTGGCGATGGACGCACAGCGC  
300 ArgLeuSerAlaIleGlyProGlyGlyLeuSerArgGluArgAlaGlyLeuGluValArg  
CGACTGTGGCTGGGGCCGGCTGTGTCACGTGAGCCGCTGGAGGTCCGC  
360 AspValHisProSerHistYrgIYArgMetCysProIleGluThrProGluGlyProAsn  
GACGTGCACCCGTCGGCACTXACGGCCGATCGAAACCCCTGAGGGGCCAAC  
420 IleGlyLeuIleGlySerLeuSerValTyrAlaArgValAsnProPheGlyPheIleGlu  
ATCCGGTCTGATCGCTGGCTGTCGGTGTACGGCGGTACGGTCAACCCGTTCTGGGTTTCATCGAA  
482 ThrProTyrArg  
ACGGCCGTACCGC

FIGURE 13

24 / 26

FIGURE 14(1)

**SUBSTITUTE SHEET**

25 / 26

FIGURE 14(2)

**SUBSTITUTE SHEET**

26 / 26

10  
R K G R R D K I G K V K T A A L K  
CGCAAGGGTC GTCGAGACAA GATTGGCAAG GTCAAGACCG CGGCTCTGAA  
10 15 20 25

30 35 40

42  
G N P Q R R G V C T R V Y T S T  
GGGCAGCCCCG CAGCGTCGTG GTGTATGCAC CCGCGTGTAC ACCACCACTC  
P K K P N S A L R K V A R V K L T  
CGAAGAAGCC GAACTCGGCG CTTCGGAAGG TTGCCCCGCGT GAAAGTTGACG  
45 50 55

60 65 70 75

E H S M V L V R G G R V K D L P  
GGAGCACTCG ATGGTGCTGG TGCAGCGGGCGG CCGGGTGAAG GACCTGCCTG  
80 85 90

95  
G V R Y K  
GTGTGCGCTAC AAG.

FIG. 15

SUBSTITUTE SHEET

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 93/01063

I. CLASSIFICATION OF SUBJECT MATTER (If several classification symbols apply, indicate all)<sup>6</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.C1. 5 C12Q1/68; // (C12Q1/68, C12R1:32)

## II. FIELDS SEARCHED

Minimum Documentation Searched<sup>7</sup>

Classification System	Classification Symbols
Int.C1. 5	C12Q

Documentation Searched other than Minimum Documentation to the Extent that such Documents are included in the Fields Searched<sup>8</sup>III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>9</sup>

Category <sup>10</sup>	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
X	JOURNAL OF MOLECULAR BIOLOGY vol. 202, no. 1, August 1988, LONDON pages 45 - 58 JIN ET AL. 'Mapping and sequencing of mutations in the Escherichia coli rpoB gene that lead to rifampicin resistance' see the whole document	1, 16, 17, 19
A	---	18, 20
Y	WO,A,9 106 674 (SCOTGEN LTD) 16 May 1991 see page 3 - page 5	1, 16, 17, 20
A	---	3-8, 18, 19
Y	EP,A,0 223 156 (HOECHST JAPAN LIMITED) 27 May 1987 see page 7 - page 10	1, 16, 17, 20
	---	-/-

\* Special categories of cited documents :<sup>10</sup>

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

## IV. CERTIFICATION

Date of the Actual Completion of the International Search

20 AUGUST 1993

Date of Mailing of this International Search Report

14.09.93

International Searching Authority

EUROPEAN PATENT OFFICE

Signature of Authorized Officer

CEDER O.

## III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
A	<p>THE JOURNAL OF GENERAL MICROBIOLOGY      vol. 60, no. 1, January 1970, LONDON      pages 125 - 132      SRIPRAKASH ET AL. 'Isoniazid-resistant      mutants of <i>Mycobacterium tuberculosis</i>  <i>H37RV</i>: ...      cited in the application      see abstract</p> <p>—</p>	2
P,X	<p>NATURE      vol. 358, no. 6387, 13 August 1992, LONDON      pages 591 - 593      ZHANG ET AL. 'The catalase-peroxidase gene      and isoniazid resistance of <i>Mycobacterium</i>  <i>tuberculosis</i>'      cited in the application      see the whole document</p> <p>—</p>	1-18
P,X	<p>THE LANCET      vol. 341, no. 8846, 13 March 1993, LONDON      pages 647 - 650      TELENTI ET AL. 'Detection of      rifampicin-resistance mutations in  <i>mycobacterium tuberculosis</i>'      see the whole document</p> <p>—</p>	1,16,17, 19
P,A	<p>RESEARCH IN MICROBIOLOGY      vol. 143, no. 7, September 1992, AMSTERDAM      pages 721 - 730      HEYM ET AL. 'Isolation and      characterization of isoniazid-resistant      mutants of <i>Mycobacterium smegmatis</i> and <i>M.</i>  <i>aurum</i>'      cited in the application</p> <p>—</p>	

ANNEX TO THE INTERNATIONAL SEARCH REPORT  
ON INTERNATIONAL PATENT APPLICATION NO.

EP 9301063  
SA 74177

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 20/08/93

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9106674	16-05-91	None		
EP-A-0223156	27-05-87	AU-B- AU-A- JP-A-	587655 6504286 62201584	24-08-89 14-05-87 05-09-87